volumeVI ♦ **issue**1
2015

Logos & Episteme

an international journal
of epistemology

Romanian Academy
Iasi Branch
◆◆◆
"Gheorghe Zane" Institute
for Economic and Social
Research

Contact address:

Institutul de Cercetări
Economice şi Sociale „Gh.Zane”

Iaşi, str.T.Codrescu, nr.2, cod 700481
Tel/Fax: 004 0332 408922
Email: logosandepisteme@yahoo.com

www.logos-and-episteme.proiectsbc.ro

This journal was edited within “Innovation and Development in the Patterning and Representation of Knowledge Through PhD and Post-PhD Fellowships” Project, supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU 159/1.5/S/133675.

The content of these materials does not necessarily represent the official position of the European Union and the Romanian Government.

For detailed information about the other programs co-financed by the European Union, we invite you to visit www.fonduri-ue.ro

ISSN 2069-053

TABLE OF CONTENTS

RESEARCH ARTICLES

A CRITICAL EXAMINATION OF BONJOUR'S, HAACK'S, AND DANCY'S THEORY OF EMPIRICAL JUSTIFICATION

Dionysis CHRISTIAS

ABSTRACT: In this paper, we shall describe and critically evaluate four contemporary theories which attempt to solve the problem of the infinite regress of reasons: BonJour's 'impure' coherentism, BonJour's foundationalism, Haack's 'foundherentism' and Dancy's pure coherentism. These theories are initially put forward as theories about the justification of our empirical beliefs; however, in fact they also attempt to provide a successful response to the question of their own 'metajustification.' Yet, it will be argued that 1) none of the examined theories is successful as a theory of justification of our empirical beliefs, and that 2) they also fall short of being adequate theories of metajustification. It will be further suggested that the failure of these views on justification is not coincidental, but is actually a consequence of deeper and tacitly held problematic epistemological assumptions (namely, the requirements of justificatory generality and epistemic priority), whose acceptance paves the way towards a generalized scepticism about empirical justification.

KEYWORDS: Laurence BonJour, Susan Haack, Jonathan Dancy, empirical justification, epistemic priority requirement, justificatory generality requirement, scepticism

1. Introduction

Most of our empirical beliefs seem at first sight perfectly justified. For example, ordinary observational beliefs (of the form "the table on which I'm writing is red" or "the chair on which I'm sitting is blue"), mnemonic beliefs ("I was watching television in the morning"), testimony beliefs ("the first world war begun in 1914") and even non-observational, scientific beliefs ("protons consist of quarks") seem to be paradigms of justified empirical beliefs. How can it be then that the justification of our empirical beliefs is considered to be a philosophical problem, which, moreover, seems to be intractable?

This is because since time immemorial[1] a sceptical argument had appeared (nowadays known as the 'infinite regress of justification' argument) which questioned the justification of beliefs *in general*, i.e. of each and every one of them, of every group or type of belief, including, of course, empirical beliefs. Its

[1] See e.g. Aristotle, *Posterior Analytics* (Oxford: Clarendon Press, 1994), 4-6, and Sextus Empiricus, *Outlines of Pyrrhonism* (Cambridge Mass: Harvard University Press, 1933), I.164-69.

basic premise is that a belief *p* is justified only if it is inferred from another belief *q* and only if *q* is an *independently justified* premise which serves as a reason or evidence in an argument with *p* -the belief under justification- as its conclusion. If this premise is granted, then the justification of beliefs in general seems to end up in an infinite regress, where a further belief is always needed for the justification of any belief whatsoever. Moreover, two familiar ways of attempting to end the infinite regress, namely, 1) circular justification where eventually – somewhere in the justificatory chain – a belief is used as a reason for the justification of itself (for example, belief *A* is justified by *B*, *B* is justified by *C*, and *C* by *A* again), and 2) arbitrary assumption, where the chain of justification ends up in regress-stopping beliefs which supposedly have the ability to justify other beliefs without being themselves justified at all, obviously fail to satisfy the above basic premise of the infinite regress argument. If we take it that the possible sources of justification ought to be completely independent of the beliefs under justification in order to be able to transmit positive epistemic status to the latter, then the only 'positive' solution to our conundrum is to posit the existence of certain 'basic epistemic units' or 'epistemic prime movers' whose justification in not inferentially transmitted to them by other beliefs. (As we shall see below, this is a description of the foundationalist position.)

In this paper, we shall describe and critically evaluate four contemporary theories which attempt to solve the above-mentioned problem of the infinite regress of reasons. Two of them, namely those of Laurence BonJour[2] and Jonathan Dancy[3] are coherentist theories of empirical justification, which reject the view that there is an absolute foundation which functions as an 'epistemic prime mover' of an essential hierarchical structure in our system of empirical beliefs, and hold instead that empirical justification is conducted within a network of mutually supporting beliefs, where some of them are more 'central' and others more 'peripheral' within the network, but in which all beliefs, central and peripheral alike, are justified solely on the basis of their contribution in the system's coherence as a whole. Next, we shall be concerned with a foundationalist theory of empirical justification, than of the 'late' BonJour,[4] according to which our system of

[2] Laurence BonJour, *The Structure of Empirical Knowledge* (Cambridge Mass: Harvard University Press, 1985).

[3] Jonathan Dancy, *An Introduction to Contemporary Epistemology* (Oxford: Blackwell, 1985).

[4] Laurence BonJour, "Foundationalism and the External World," *Philosophical Perspectives* 13 (1999): 229-49; Laurence BonJour, *Epistemology: Classic Problems and Contemporary Responses* (Maryland: Rowland & Littlefield, 2002); Laurence Bonjour and Ernest Sosa, *Epistemic Justification, Internalism vs. Externalism, Foundations vs. Virtues* (Oxford: Blackwell, 2003).

empirical beliefs ultimately rests on non-conceptual yet conscious and epistemically efficacious sensory experience; the latter provide the epistemic foundation on which the whole edifice of empirical knowledge is built (by inference from sensory experiential content). And, lastly, we will examine a 'hybrid' theory, that of Susan Haack,[5] according to which the correct view about empirical justification is one which combines foundationalist and coherentist elements while at the same time rejecting certain problematic features that equally occur to these two opposing views.

As was mentioned above, those theories are initially put forward as theories about the justification of our empirical beliefs; however, in fact they also attempt to provide a successful response to the question of their own '*meta*justification' ("do I have any reasons to believe that the criteria of justification provided by my own theory of justification are *true*?"). Hence, it seems that the above theories also attempt to connect empirical justification with empirical knowledge in a non-sceptical manner. Yet, it will be argued that 1) none of the examined theories is successful as a theory of justification of our empirical beliefs, and that 2) they also fall short of being adequate theories of metajustification. That is to say, even if they indeed were adequate as theories of empirical justification, they would fail to provide a non-sceptical connection between empirical justification and empirical *truth* (or empirical *knowledge* – since the truth of our beliefs is a necessary condition for them to constitute knowledge).

After a detailed presentation of BonJour's, Haack's and Dancy's theory we shall attempt to show, mostly by means on internal critique, that they cannot avoid ending up to certain sceptical conclusions regarding empirical justification. That is to say, they do not succeed in solving an epistemological problem which all of them admit as legitimate, namely that of their own 'metajustification.' It will be further suggested that the failure of these views on justification is not coincidental, but is actually a consequence of deeper and tacitly held problematic epistemological assumptions, whose acceptance paves the way towards a generalized scepticism about empirical justification.

2. BonJour's Coherentist Theory of Empirical Justification

BonJour's coherentist theory of justification is presented in his well-known book *The Structure of Empirical Knowledge*. The basic ideas around which his theory is structured are the following: 1) The holistic nature of empirical justification (the unit of justification is the whole system of beliefs, rather than individual beliefs),

[5] Susan Haack, *Evidence and Inquiry* (Oxford: Blackwell, 1993); Susan Haack, "Reply to BonJour," *Synthese* 112 (1999): 25-35.

2) a clarification of the concept of coherence, 3) the idea of 'doxastic presumption' (can I grasp my own belief system, and if yes, to what extent?), and 4) the idea of the observational requirement, which is BonJour's way of accommodating observation within a resolutely coherentist framework.

According to BonJour, the holistic coherentist model of empirical justification is precisely what is needed to solve the problem of the infinite regress of reasons. This is so because, according to the coherentist conception, justification is not 'transferred' from one belief which is already justified to another which is 'waiting' to receive justification, through inferential relations (of epistemic priority), but is instead an epistemically 'simultaneous' property of the *whole system* of beliefs, namely the property of the system's coherence as a whole. Only if one held the former 'atomistic,' non-holistic view of empirical justification, which coherentism explicitly and resolutely *rejects*, would it follow what many epistemologist take for granted without much further argument, namely that coherentist justification is viciously circular.

Now, BonJour explicates the crucial concept of 'coherence' as follows: 1) A conceptual system is coherent only if it is logically consistent (although he eventually changed his mind and came to the view that a) logical consistency need not be an absolutely necessary condition of a system's coherence, and b) a system of beliefs can well have a high degree of coherence in spite of the presence of 'local' incoherence within it.[6] 2) The coherence of a system of beliefs is proportional to the degree of its probabilistic consistency, which can be understood as the demand that the system, besides the belief *p* ought also to include the belief that *p* is *probable*. 3) The coherence of a system of beliefs is increased if those beliefs are inferentially related, and this increase is directly proportional to the number and strength of the inferential relations in question. 4) The coherence of a belief-system is decreased if there exist within it several subsystems of beliefs which are relatively unconnected (i.e. inferentially unrelated). Only inferential relations can provide mutual epistemic support among beliefs. Two or more beliefs are inferentially related if one can serve as a premise of an argument for the justification of another. 5) However, the mere fact that certain beliefs are inferentially related is not sufficient for obtaining ideal coherence since it does not exclude the possibility that two or more conceptual systems may be well be coherent considered in isolation from one another, i.e. without being *themselves* inferentially connected. Intuitively, it seems that an ideally coherent system is one in which its relatively independent subsystems are

[6] Laurence BonJour, "Replies and Clarifications," in *The Current State of the Coherence Theory*, ed. John Bender (Dordrecht: Kluwer, 1988), 284.

inferentially connected by higher level explanatory laws and principles. 6) The coherence of a belief-system is decreased to a degree which is directly proportional to the presence of unexplained anomalies within it.[7] This means that for a proper definition of the concept of coherence we must take into account the existence of *explanatory* relations among beliefs (over and above their mere inferential connections).

As was mentioned above, a further important element of BonJour's coherentist theory is the doxastic presumption. According to it, our beliefs about the contents of our own beliefs, i.e. our conception of our own system of beliefs, is for the most part true. This presumption cannot itself be justified by appeal to its relations of coherence with other beliefs of the system since that would presuppose what is in question, namely the truth of one's (meta)belief about the contents of one's own beliefs. Our epistemic practices can get off the ground only if they already take it for granted that one's beliefs about the contents of one's own beliefs are for the most part true; otherwise our epistemic practices would lose their point, and the very attempt to raise a question about the justification of any of our beliefs would be pointless.

Next, BonJour attempts to provide a plausible response to one of the most pressing questions that every coherence theory sooner or later has to face, namely that of making room for a plausible concept of observation within the confines of a coherence theory. How can observation function as an independent epistemic check of other, non-observational, empirical beliefs of the system in the context of a generalized coherence theory of empirical beliefs? Absent this crucial epistemic function of observation, it seems that there is nothing in a coherence theory of justification which necessitates that the content of empirical beliefs within the system is ultimately provided by the external world, i.e. a world that exists and has certain structural and qualitative properties independently of their representation from within our system of beliefs. BonJour's response to this problem is that, provided that we accept the doxastic presumption, it is possible to identify a sub-class of beliefs, namely that of *cognitive spontaneous* beliefs (i.e. beliefs which are non-inferential in origin) and to infer that *some* kinds of those cognitive spontaneous beliefs – e.g. introspective or observational ones, as contrasted with cognitively spontaneous beliefs that are the result of wishful thinking, 'hunches' or unfounded irrational dogmas- are, from the standpoint of the system of beliefs of which they are part, highly reliable, therefore probably

[7] A conceptual system is plagued with theoretical anomalies if some of its beliefs entail the existence of a state of affairs which cannot be explained on the basis of other beliefs of the system.

true. The belief in the reliability of the above groups of cognitively spontaneous beliefs is based on the fact that the individual members of the group are in agreement with one another as well as with members of other groups of spontaneous beliefs. And this 'agreement' consists in the absence of anomalies between them, in their 'hanging together' with other theoretical principles which contribute to the formulation of an ever-increasing coherent picture of an independent, objective reality, and, crucially, in the existence of a hypothesis which *explains* their reliability. Cognitively spontaneous beliefs which fulfil the above requirements can justifiably be considered as *observational* – while the non-observational beliefs of the system can only be justified if they cohere with precisely those observational beliefs.[8]

However, BonJour himself observes that at this stage of his argument his coherence theory allows only for the *possibility* of external, 'independent' input to the 'internal' process of coherentist justification of our empirical beliefs; it does not guarantee that this is actually the case. This is because it is possible that a system of beliefs which entails that certain recognisable kinds of cognitively spontaneous beliefs are very likely to be true can nonetheless *fail* to imply that the content of those beliefs is reliably correlated to external worldly causes. Therefore, a coherentist theory which purports to be an essentially (and not just accidentally) reliable guide to empirical truth ought to require that the individual beliefs of the system can be candidates for empirical justification only if the system includes laws which ascribe a high degree of reliability to an extended variety of cognitively spontaneous beliefs. In BonJour's terminology, this is the 'observational requirement,' which is necessary for the viability of any coherence theory of empirical justification.[9]

However, BonJour's coherence theory is not without its problems. Its major problematic elements (which, interestingly, remain as such even in his 'late' foundationalist turn) stem from his deep-seated commitment to a *strong internalism* combined with an *argumentative/inferential model* of understanding empirical justification. Those epistemological commitments seem to seriously undermine his theory since they seem to imply that empirical justification is actually a process without end. This is because, according to BonJour's theory, one's empirical belief (whatever its content may be) can be justified only if one can explicitly grasp and justify inferentially (from prior justified beliefs), first, its own content (i.e. the fact that its content is what the subject believes it to be), and, second, its relations of coherence with the contents of the beliefs with which it is

[8] BonJour, *Coherence Theory*, 138.

[9] BonJour, *Coherence Theory*, 141.

inferentially connected. To his credit, BonJour does not ignore this potentially devastating sceptical consequence of his commitment to internalism, and he attempts to deal with the problem by introducing his 'doxastic presumption.' Does this move solve the problem of epistemological scepticism?

A direct consequence of the above-mentioned combination of internalist-argumentative model of justification is that the justification of a cognitively spontaneous belief (however obvious its content may seem to us) *presupposes* an epistemically prior justification of the (meta)belief to the effect that the content (and origin) of the cognitively spontaneous belief in question is what the subject thinks it is and not something else. But this latter (meta)belief can be justified only if the doxastic presumption is *itself already justified* (and not just *true*) since, according to the above internalist-argumentative model of justification, an empirical belief can be justified only if one has a good reason for believing it, and a reason for believing something is a good – i.e. non question-begging – reason only if it is epistemically prior to the belief under justification. BonJour himself accepts that the doxastic presumption cannot be itself justified since its truth is a necessary condition of the possibility of the justification of any belief whatsoever. But, from this it follows either that 1) the part of the justification of a spontaneous belief which depends on the doxastic presumption is conferred to this belief only from the fact (if it is a fact) that the doxastic presumption *happens* to be true (not from the fact that we have *good reasons* to believe that it is true), or that 2) since the doxastic presumption is not itself justified, then neither are our spontaneous beliefs about the contents of our beliefs nor our further, spontaneous or not, first-level empirical beliefs of ours (since the justification of the latter is conditional on the justification of the former).

If we take the first option, it follows that this way of justifying the doxastic presumption is not internalist at all since the subject does not possess good (non question-begging) reasons to believe that the doxastic presumption is true. The only reason to believe such a thing would be that if the doxastic presumption were false, the enterprise of empirical knowledge could not even get off the ground. Yet, this kind of 'reason' does not seem to be genuinely epistemic by Bonjour's own lights (since it presupposes that there actually is empirical knowledge), but is rather an expression of a deep human 'desire' for empirical knowledge or of a practical inability to conceive the epistemic possibility of the non-existence of any empirical knowledge whatsoever. If, on the other hand, we take the second option, then, given Bonjour's commitment to a strong internalism combined with an argumentative-inferential model of justification, we end up to a radical scepticism regarding the justification of each and every empirical belief of ours.

3. BonJour's Foundationalist Theory of Justification

BonJour eventually came to the conclusion that the above critique regarding the epistemic status of the doxastic presumption is eminently plausible, and acknowledged that, given his commitment to a strong internalist and argumentative-inferential model of empirical justification, only a *foundationalist* interpretation of this presumption could block the way to a radical scepticism. Accordingly, in a series of articles from 1998 to 2003[10] he ended up advocating a pure foundationalist theory of empirical justification in a more or less traditional form of this view which has its origins in Locke's and Berkeley's empiricism. BonJour thereby attempted to resuscitate a version of the traditional epistemological concept of the Given, as the latter was used for providing a foundation to empirical knowledge in the works of Russell[11] and C.I. Lewis.[12]

According to this new, foundationalist theory of justification, a certain subset of our empirical beliefs, namely that of 'basic' beliefs, are non-inferentially justified, i.e. justified in a way which automatically makes them intrinsically reliable independently of any epistemic support from other beliefs. This is possible because the justification of basic beliefs is internally related to the (non-conceptual) content of sensory and introspective experience, and, more specifically, to the *immediate apprehension* of the content of sensory experience.[13]

Now, according to BonJour, there are two kinds of basic beliefs: Meta-beliefs about the content of first-order beliefs (notice that these are precisely those meta-beliefs which constitute BonJour's 'doxastic presumption,' which are now foundationally justified), and beliefs about sensory experience. The latter purport to (conceptually) describe – in physical object appearance terms ("it looks as if there is a blue book in front of me") – the non-conceptual, 'phenomenal' content of sensory experience, and they are true if they provide an *accurate description* of that content. This description can be correct or incorrect (unlike the immediate apprehension of sensory experience which is neither correct nor incorrect), but the subject is in an ideal epistemic position to grasp the correctness of incorrectness of the description due to the fact that the content to be described is

[10] Laurence BonJour, "The Dialectic of Foundationalism and Coherentism," in *The Blackwell Guide to Epistemology*, eds. John Greco and Ernest Sosa (London: Blackwell, 1998), 117-44; BonJour, "Foundationalism and the External World"; BonJour, *Epistemology*; Bonjour and Sosa, *Epistemic Justification*.

[11] Bertrand Russell, *The Problems of Philosophy* (Oxford: Oxford University Press, 1969).

[12] Clarence Irving Lewis, *Mind and the World* Order (New York: Dover, 1929); Clarence Irving Lewis, *An Analysis of Knowledge and Valuation* (LaSalle, IL: Open Court, 1946).

[13] BonJour, *Epistemology*, 63-64.

itself a *conscious* mental state of the subject. BonJour acknowledges that there is always the possibility of error in descriptions of this kind (empirical justification is defeasible) but, precisely because in this juxtaposition of the conceptual content of the description with the non-conceptual content of sensory experience the subject is in an ideal epistemic position to judge the accuracy of this description, empirical justification at this level can lose its *prima facie* reliability only if there is a special, specific reason for one to believe that the description in question is inaccurate (e.g. distraction, inattention etc.); the abstract possibility of the existence of such a defeating reason is not sufficient for putting the *prima facie* reliability of basic beliefs into question.[14]

Unfortunately, BonJour's foundationalist theory of justification faces serious problems, in spite of the fact that (or rather, as we shall see in section 6, *precisely because*) it is diametrically opposed to his former coherentist view. More specifically, it is unclear how, on this view, empirical justification is generated in the case of basic beliefs.

As regards the question of the generation of the empirical justification of basic beliefs, the problem in BonJour's foundationalism is that it goes hand in hand with a strong form of epistemological *Givenness*, according to which the self-presenting properties of the act of immediate apprehension of the content of sensory experience are sufficient to guarantee that the conceptually structured basic belief which purports to describe that non-conceptual, sensory content are intrinsically *prima facie* justified. Yet, in our view, BonJour's own understanding of those 'self-presenting' properties of the non-conceptual act of immediate apprehension is such that it deprives another related, and most fundamental, epistemic act, that of the *direct comparison or juxtaposition* of the conceptual and the non-conceptual content of perceptual experience, of its very *normativity*, i.e. of a necessary condition for a basic belief be considered as *prima facie* justified (or unjustified) at all. This latter epistemic act of 'direct comparison/juxtaposition' can be considered as capable of being correct or incorrect (justified or unjustified) only if there are *independent criteria* of distinguishing between cases in which it *seems* that this 'direct comparison' has the content we think it has (while in reality, its actual content is not what we think it is) and cases in which it *actually has* the content we think it has.[15] This absolutely crucial distinction for the viability of BonJour's whole foundationalist epistemological project cannot be drawn in the BonJourean allegedly 'epistemic' act of 'direct comparison' since the very content

[14] BonJour, *Epistemology*, 213-15; BonJour and Sosa, *Epistemic Justification*, 72-75.

[15] See also Wilfrid Sellars, "The Structure of Knowledge," in *Action, Knowledge and Reality*, ed. Hector-Neri Castaneda (Indianapolis: Bobbs-Merill, 1975), §24-25.

of the latter is such (phenomenal) that the notion of an independent identification of its correct or incorrect recognition and application cannot find any foothold.

If, on the other hand, as BonJour suggests in a reply to his critics,[16] the epistemic act of 'direct comparison' is in fact partly conceptual and clearly distinct from the contents under comparison, namely those of the perceptual belief and the non-conceptual sensory experience, then it seems that, ironically, BonJour falls prey to a critique that he himself levelled against foundationalism in his coherentist period. If one introduces semi-judgemental or semi-conceptual states in an attempt to provide an independent epistemic criterion or foundation for the justification of empirical beliefs one just does not recognise the plain fact that "to whatever extent such a state is capable of conferring justification, it will to that very same extent be itself in need of justification."[17,18]

It seems therefore that the foundationalist idea of a level of intrinsically justified basic beliefs is highly problematic in the radical sense that it cannot even get off the ground.

4. Haack's 'Foundherentist' Theory of Empirical Justification

It seems that BonJour's foundationalism fares no better than his older coherentist theory in solving the problem of empirical justification. But there are at least two further theoretical positions in the vicinity which could be worth pursuing: The first is the return to coherentism, albeit of a different form than BonJour's former 'impure' coherentist theory (such as Dancy's 'pure' coherentist theory, which will be examined in section 5) while the second is that of some kind of synthesis or 'fusion' of foundationalism and coherentism. In this section, we will discuss Susan Haack's attempt to provide such a theoretical fusion in her book *Evidence and Inquiry.*[19]

Haack terms her theory 'foundherentism' in order to stress that this view incorporates central elements from both foundationalism and coherentism. She further believes that although those theories are diametrically opposed to one another, they are not absolutely dichotomous, that is, they do not exhaust the

[16] Laurence BonJour, "Reply to my Critics," *Philosophical Studies* 131 (2006): 743-59.
[17] BonJour, *Coherence Theory*, 78.
[18] For further, but in many ways similar arguments against BonJour's foundationalism (which, moreover, point out that an adherence to the doctrine of the Given conflicts with the requirements of epistemological internalism) see Byeong Lee, "BonJour's Way Out of the Sellarsian Dillemma and his Explanatory Account," *Dialogue* 52, 2 (2013): 287-304, and Ted Poston, "BonJour and the Myth of the Given," *Res Philosophica* 90, 2 (2013): 185-201.
[19] Haack, *Evidence and Inquiry* .

theoretical space of possible philosophical responses to the central problems of epistemology (such as the 'infinite regress argument').

According to Haack, the primary source of empirical justification is evidence, which can be either propositional or non-propositional in character. That is, empirical justification is not only a matter of having reasons to believe something – since, according to Haack, reasons can be only propositional in form.[20] For example, sensory experiences of which we are aware about what *seems* to be the case (sensory 'seemings') can well function as non-propositional evidence for an empirical belief. Haack terms those sensory seemings "S (state)-evidence (as contrasted to 'C (content)-evidence" which are explicitly propositional and function as reasons). Indeed, it seems that the notion of 'S-evidence' can provide the key for a proper response to the problem of the infinite regress of justification since S-evidence are in part epistemically efficacious (i.e. they can justify other beliefs) without themselves being further beliefs (in need of further justification).

Haack's theory combines elements of foundationalism and coherentism without being reduced to either of them. On the one hand, she accepts the foundationalist view that purely causal factors can make a contribution to the justification of an empirical belief, while, on the other hand, she contends that the justification of our empirical beliefs *as a whole*, including basic beliefs, cannot be adequate unless at least some conceptual evidence, i.e. evidence that belong in the 'space of reasons,' are used as premises in a justificatory argument. In this sense, Haack is not a foundationalist since she rejects the epistemic significance of the divide between basic and non-basic beliefs.[21]

In another conciliatory move, Haack also purports to show that the dichotomy between reasons and causes – exactly like that between foundationalism and coherentism – is not absolute. She does that by maintaining that certain kinds of evidence within the space of reasons, namely 'experiential C-evidence,' can adequately capture the justificatory power that certain kinds of non-conceptual states situated within the 'space of causes' (experiential *S*-evidence) are supposed to possess. If this justificatory connection between the conceptual and the non-conceptual level could indeed be demonstrated, Haack's further claim to the effect that propositions which describe experiential C-evidence are all true (i.e. propositions of the form "I undergo a sensory or perceptual experience of a such-and-such kind") would be eminently plausible; and in this way, Haack could claim that her theory can successfully anchor our conceptual system of empirical beliefs to the non-conceptual structure of sensory

[20] Haack, *Evidence and Inquiry*, 70.

[21] Haack, *Evidence and Inquiry*, 19.

experience, and eventually, to the extra-conceptual structure of the external world. Haack suggests that the way in which experiential C-evidence can capture the justificatory relevance of experiential S-evidence is by (conceptually) ascribing or positing the occurrence – under normal conditions – of a specific (non-conceptual) sensory experience which has definite subjective experiential features, but whose content can be described in terms of external physical objects and events.[22].

However, at this point one can raise the following objection: What reason do we have to believe that the conceptual description of experience can capture exactly those properties of our non-conceptual experience which are relevant to its justificatory power? How do we know that our general beliefs with regard to the ultimate causes of the occurrence or the structural and qualitative properties of our non-conceptual sensory experience are justified?[23] Haack's responds by invoking our common sense concept of evidence and by noting that at this stage of her argument she only purports to describe our pre-theoretical criteria of empirical justification, rather than evaluating them as regards their capacity to be reliable indications of empirical truth. (This latter project is that of the 'ratification' of our existing criteria of empirical justification.) Haack therefore suggests that a successful response to the abovementioned serious BonJourean objection goes hand in hand with the provision of an illuminating answer to the further question of the ratification of her own (purely descriptive, commonsensical) criteria of justification.[24] But before we examine Haack's attempt to solve the 'metajustification' problem, we should first get a clearer (if only sketchy) grasp of her first-level 'descriptive' or 'commonsense' theory of empirical justification.

Can we provide a plausible account of our pre-theoretical criteria of justification? What conditions must be met in order for our evidence to count as *good* evidence for believing something about the empirical world? We can better understand what these condition are if we use as our analogy the model of the *crossword puzzle*. The clues of the crossword puzzle correspond to one's experiential C-evidence while the already filled-in entries correspond to the C-reasons one has for believing something about the world. The clues that are initially given in the crossword puzzle do not depend on its already filled-in entries, while the latter's accuracy does depend on the accuracy of the other filled-in entries. In exactly the same sense, the justification of experiential C-evidence

[22] Haack, *Evidence and Inquiry*, 80-81.
[23] Laurence BonJour, "Haack on Justification and Experience," *Synthese* 112 (1997): 13-23.
[24] Haack, *Evidence and Inquiry*, 17.

for believing something does not depend on the C-reasons that one may possess for it, but the justification of the C-reasons themselves does depend on the justification of other C-reasons we have for that belief.[25] The clues of the crossword puzzle correspond to our conceptual responses to our sensory experience (as the former are formulated in propositions about our own perceptual experience) which ultimately results from the causal triggering of non-conceptual informational states of the (pre- or extra-subjective) world; the already filled-in entries of the crossword puzzle correspond to the conceptually structured reasons which, by their coherence relations to each other provide justificatory support to our empirical beliefs, while the current status of the process of puzzle-solving, i.e. the degree in which it has been filled-in and the correctness or incorrectness of the completed entries correspond to the current status of our conceptual system of empirical beliefs with respect to the degree in which it accurately represents the external world.

Now, based on the above analogy of the crossword puzzle, Haack suggests that the justification of empirical beliefs ultimately depends on 1) the degree of supportiveness which is given to an 'entry' (belief) by the initial 'clue' (S-evidence) as well as from any intersecting entries that have been already filled in (reasons), 2) how reasonable, *independently* of the entry in question, one's confidence is that those other already filled-in entries are correct, and 3) how many of the intersecting entries have actually been filled in.[26]

But how are Haack's criteria of empirical justification *themselves* justified? Do we have any reason to believe that those criteria are a reliable guide to empirical truth? As was mentioned above, this is the project or ratification or metajustification of the criteria of justification. Haack herself describes the most central and general difficulty that any theory of justification is bound to face as regards its own metajustification as follows: Let 'R' abbreviate all the direct reasons that can be offered in a ratificatory argument, from the standpoint of a 'foundherentist' (or, for that matter, any) theory of justification. Even then, there remains the question 'how do we know that R?' To this question one can respond either by begging the question – i.e. by using what from the standpoint of the theory in question (Haack's foundherentism) are standards of good (i.e. truth-indicative) evidence and then showing that the foundherentist standards for R satisfy the standard of good evidence – or one can use an externalist argument to the effect that even if one does not know that one's evidence for R are indeed good (i.e. supportive and independently secure) one can nonetheless be justified in

[25] Haack, *Evidence and Inquiry*, 81-82.

[26] Haack, *Evidence and Inquiry*, 82.

believing them since they may well *be* good evidence. Haack opts for the externalist response to the ratificatory question.[27]

However, it seems that Haack's externalist response to the ratificatory question cannot be satisfactory since it is patently question-begging even from the standpoint of her own first-order 'foundherentist' criteria of empirical justification. Specifically, Haack's externalist argument to the effect that the only thing that is sufficient for one to know that one's belief about R is justified is just the actual truth of one's belief about R (rather than its independent justification) obviously violates her own (first-order) justificatory standard of *independent security*.

Another serious objection against Haack's foundherentism, which is targeted against its own conceptual core, is the following: As was shown above, Haack suggests that there is a middle ground between foundationalism and coherentism, in which the justification of *every* empirical belief essentially involves doxastic as well as non-doaxastic sources of justification. Yet, as it appears to be the case, Haack holds that no empirical belief whatsoever can be epistemically efficacious (i.e. transmit justification to other beliefs) unless it is justified from *both* sources of justification, it follows that the doxastic sources of justification cannot be usefully distinguished from non-doxastic ones with respect to their capacity to render an empirical belief epistemically *efficacious*. Hence, it seems that Haack's insistence that the non-doxastic factors of justification provide a 'positive epistemic status' to an empirical belief all by themselves (even if only in part) is not adequately motivated since she also explicitly holds that this 'partial' justification is never sufficient for making a belief capable of being epistemically efficacious. Why should we take it that non-doxastic sources of justification provide positive epistemic status to an empirical belief all by themselves (rather than, say, always with the aid of doxastic factors) if this positive epistemic status does not even suffice for enabling it to be epistemically efficacious?

This objection to Haack's foundherentism seems even more pressing if we consider Haack's own response to BonJour, in which she takes up the exact same question we raised above about the proper way to distinguish basic from non-basic beliefs in very weak forms of foundationalism (where this distinction seems to be blurred).[28] Haack correctly observes that if a version of weak foundationalism allows for a form of justificatory dependence of basic beliefs on non-basic beliefs which is so pervasive that without this support basic beliefs cannot justify other beliefs (i.e. they are not epistemically efficacious), then why not allow this

[27] Haack, *Evidence and Inquiry*, 221.

[28] Susan Haack, "Reply to BonJour," *Synthese* 112 (1997): 25-35.

pervasive justificatory support to extend to the initial 'positive epistemic status' which basic beliefs allegedly receive solely from non-conceptual sensory experience? Any prohibition of this extension cannot but be totally *ad hoc*, artificial, and, ultimately, epistemically arbitrary since it does not seem to explain anything at all, beyond our desire to continue to call our theoretical position 'foundationalism.' But exactly the same objection can be levelled against Haack as regards the necessity (epistemic function) of her distinction between doxastic and non-doxastic sources of empirical justification. Haack believes that this epistemic distinction ought to be preserved at all costs because only in this way can we solve a problem which is fatal for all forms of coherentism, namely that of accounting for the possibility of epistemically representing the external, extra-conceptual world from the inside of our conceptual system of beliefs and solely on the basis of their internal epistemic relations of coherence. Yet, in our view, Haack's foundherentism does not succeed in solving this problem, save only verbally. This is because Haack's non-doxastic sources of justification (experiential S-evidence) – which supposedly ensure that our conceptual system is epistemically anchored to external, extra-conceptual reality by providing an epistemic source outside beliefs that can function as an independent check of the latter- are in fact epistemically *idle* since they cannot even transmit their supposed 'partial' justificatory power (their 'positive epistemic status) to our beliefs (i.e. to the conceptual level) without justificatory recourse to the *doxastic* or conceptual level. Hence, they cannot play the role (not even in part) of an extra-conceptual level, outside all beliefs, which can provide a (partly) independent epistemic check of our conceptual system of empirical beliefs.

5. Dancy's Pure Coherentist Theory of Justification

It starts to seem that neither BonJour's coherentism or foundationalism, nor Haack's 'hybrid' 'foundherentist' theory can provide a viable solution to the problem of empirical justification. Haack's foundherentism seems on careful inspection to collapse to a kind of coherentism, and, as was shown in section 3 BonJour's foundationalism cannot even get off the ground (unless it appeals to essentially coherentist considerations). Hence, it seems reasonable to suppose that a coherence theory of empirical justification which could avoid the pitfalls of BonJour's version of coherentism could stand a chance of being a correct view about empirical justification. To this end, we shall proceed to an examination of

Dancy's pure coherence theory, as the latter is developed in his book *An Introduction to Contemporary Epistemology*.[29]

According to Dancy, a belief *p* is justified to the extent to which it is a member of a coherent system of beliefs. That is, if the coherence of one's system of beliefs would be increased by abandoning the belief and by replacing it by its opposite, the belief is not justified, whereas if one's system of beliefs is more coherent with the belief that *p* as a member rather than with any alternative, the belief is justified. The basic feature of this coherentist view of justification is that it rejects the view that there is an important (i.e. asymmetric) epistemic distinction to be made between basic (non-inferentially justified) and non-basic (inferentially justified) beliefs. *Each* of the beliefs which are members of a coherent system is justified in an epistemically '*simultaneous*' manner, as it were (which is not to say that they are *inferentially* justified, in the foundationalist preferred interpretation of the term -itself essentially based on an asymmetric notion of justification), to the extent to which each belief contributes to the coherence of the system as a whole. Moreover, a system of beliefs cannot be coherent unless the contents of its beliefs are connected with relations of mutual explanation and do not produce any inconsistencies.[30]

The fact that in Dancy's version of coherentism all empirical beliefs of the system are symmetrically justified deprives the question of the metajustification of basic first-level coherentist criteria of justification[31] of any sense. It is precisely at this crucial point that Dancy's coherentism is far more 'resolute' than BonJour's 'impure' coherentism, in which the above question of metajustification is considered as perfectly legitimate and, precisely for this reason, by BonJour's own lights, it can only be answered by appeal to epistemic standards and principles which are *not themselves* coherentist in nature (think e.g. of BonJour's key non-coherentist notions of the doxastic presumption and the observational requirement). However, in this way BonJour reintroduces the (essentially foundationalist) *epistemic priority* requirement as a necessary condition of empirical justification (albeit, not on the level of individual beliefs, but at that of belief *systems*) and thereby his (supposedly coherentist) position eventually becomes foundationalist in character -which, moreover, explains why BonJour's problems with coherentism led him directly to *foundationalism*, rather than 'foundherentism' or another version of coherentism. Dancy's coherentism avoids

[29] Jonathan Dancy, *An Introduction to Contemporary Epistemology* (Oxford: Blackwell, 1985).
[30] Dancy, *Contemporary Epistemology*, 111-12, 116-18.
[31] Those first-order coherentist criteria of justification are, for example, simplicity, unification, empirical adequacy, minimization of ad hoc hypotheses or explanatory anomalies etc.

the above pitfall, but, as we shall see in what follows, it is vulnerable to an objection that can be levelled against all forms of pure coherentism (i.e. forms of coherentism in which justification is exclusively symmetrical).

Now, Dancy is not only a pure coherentist about empirical justification, but is also a coherentist about empirical *truth*. According to the latter, a belief is true if and only if is a member of a coherent system of beliefs – albeit, not of any old system of beliefs, but of an *ideally* coherent system (notice that this is a view about the criterion of truth, not a definition of truth). It does not follow that every time that there is an increase in the degree of the system's coherence the latter better approximates how the world really is, but, according to Dancy, it does follow that the increase in the system's coherence gives us a reason to believe that its beliefs are true (and Dancy thinks so because, in his view, although justification is not the same concept as truth, both justification and truth are internal relations of mutual explanation among the members of a belief set).[32] Dancy also believes that coherentism about truth can be combined with a *robust realism* about truth, namely with a sense in which truth exceeds all evidence we could possibly have for it. More specifically, Dancy suggests that an ideally coherent system of beliefs *transcends* every possible system of beliefs -however large and coherent this might be- and for every such set of beliefs there is a possibility that its beliefs are false. This results from the fact that there is a *logical, or better, epistemic gap* between the epistemic *finitude* which cannot but characterize every possible belief set and the *infinity* which is inherent in the notion of the ideally (i.e. fully comprehensive) coherent system.[33]

Yet, if Dancy's realism about truth is such that it is possible for what is true to exceed *all possible* evidence which can be offered for it from the standpoint of our epistemic practices then Dancy seems to have opened an unbridgeable epistemological gap between justification and truth, which cannot be bridged by their supposed 'internal relation.' This epistemological gap is unbridgeable because it can be opened even in the limiting case in which we consider 'all possible' evidence for believing something. Even in this epistemically ideal case, our necessary epistemic 'finitude' deprives us of the right to reasonably believe that what is internally justified is true of the external world. What reason do we have to believe that essentially 'finite' reasons can ever be reliable indications of the true nature of an 'infinite,' radically external reality? It seems that this epistemological gap is a result of Dancy's radically different conceptions of what a conceptual system is. Considered as a criterion of empirical justification, a

[32] Dancy, *Contemporary Epistemology*, 112, 116-17.
[33] Dancy, *Contemporary Epistemology*, 138-39.

conceptual system is always epistemically within our reach; yet, considered as a criterion of empirical truth, it is something that (rather mysteriously) can always be outside of our epistemic reach in the radical sense that it transcends our very epistemic 'human condition,' namely our *'constitutive' epistemic finitude*. Does it not start to seem that, ironically, Dancy's supposed 'internal relation' between empirical justification and truth ends up being pretty compatible with a correspondence theory of truth according to which the relation between our justified beliefs and states of affairs of the external world is *external* through and through?

Moreover, a major problem for the viability of Dancy's pure coherentism – which is a consequence of his basic thesis to the effect that all empirical beliefs of the system are justified in the same (i.e. symmetrical) manner – arises when one realizes that there is an inherent ambiguity in the determination of the epistemic status that the theory attaches to a special kind of beliefs in the system, namely our beliefs about *what coherence itself really is*. For example, according to those beliefs about coherence, relations of coherence are inferential relations, and more exactly, a proper subset of them: relations of mutual explanation. These, in turn, are understood in terms of general methodological criteria of simplicity (i.e. unification of individual beliefs of the system by appeal to the smallest possible number of unrelated ultimate theoretical principles ('unexplained explainers') needed for their explanation), empirical adequacy, minimization of *ad hoc* hypotheses and of the system's anomalies and so on. Now, it follows from these very methodological principles which determine the meaning of the concept of coherence that any belief of our conceptual system can be *revised*. And this means that this possibility holds equally for these methodological principles *too*. But if the coherentist criteria of justification are not themselves stable, unrevisable, and can, instead, be revised, altered or rejected, then the very *meaning* of the concept of coherence becomes wholly indeterminate for epistemological purposes. If, on the other hand, we suppose that these coherentist criteria cannot be revised, then our beliefs about them cannot any more be considered to be justified in a coherentist manner and become instead *foundationalist basic beliefs*. Thus, the core coherentist position about the symmetrical nature of the justification of all our empirical beliefs is wholly undermined and the supposedly 'coherentist' theory is transformed into a kind of foundationalism.[34]

[34] For a similar kind of argumentation against coherentism see Michael Williams, *Problems of Knowledge* (Oxford: Oxford University Press, 2001), 134-35. It should also be noted here that coherentism falls victim to the above argument only if it holds that the methodological principles about coherence are justified *in the same (i.e. symmetrical) manner* as any other

6. Is There a Sense in Which All the Above Theories of Empirical Justification Share the Same Background Problematic Epistemological Assumptions?

It seems that all the above theories of justification, whether they be coherentist (early BonJour, Dancy), foundationalist (late BonJour) or somewhere in between coherentism and foundationalism (Haack) are equally problematic. Why is that so? Is it that the problems that plague both foundationalism and coherentism originate from a common root (notwithstanding the radical opposition between these theories in a surface level), or are they unconnected?

In my view, the fact that those otherwise radically different theories of justification are all deeply problematic is not coincidental, but is rather a necessary consequence of their largely unacknowledged adherence to certain background problematic epistemological assumptions.

The most important epistemological assumption of this kind is probably the requirement of *justificatory generality*, according to which justification can be satisfactorily obtained only if it is *completely general*. Justification is completely general if a whole set of beliefs with a common subject-matter (e.g. empirical beliefs) can be justified only from an epistemic standpoint which is a position to explain how any knowledge of a set of beliefs with the subject-matter in question is possible at all. And the required justificatory generality, in turn, can be obtained (that is, vicious justificatory circularity can be avoided) only if another related epistemological requirement is satisfied, namely that of *epistemic priority*. According to the latter, a reason can justify beliefs with a common subject-matter (e.g. empirical beliefs) only if it is *itself justified* from an epistemic source whose content is *independent* of the content of the (set of) beliefs under justification.[35]

We claimed above that if one attempts to satisfy justificatory generality without accepting epistemic priority would end up in a vicious justificatory circle. But why exactly the circularity that would arise by attempting to satisfy justificatory generality without accepting epistemic priority must be vicious? To see why, consider the well-known example of the problem of justifying everyday

---

particular empirical belief of the system (that is, by their coherence relations to those particular beliefs). What then if we suppose that these methodological principles are not justified by virtue of coherence relations holding between them and the particular empirical beliefs of the system, but in some other, non-coherentist, manner? In that case, the resulting position is immune to the above objection, but, unfortunately, at the expense of not being a recognizable coherentist position at all (as was the case with BonJour's impure coherentism which turned out to be a version of foundationalism).

[35] Barry Stroud, *Understanding Human Knowledge* (Oxford: Oxford University Press, 2000), 99-101.

knowledge of the objects around us. If one asks how anyone can know anything at all about objects around us it would be obviously inadequate to reply that, for example, I know that my neighbour is at home by seeing his car in front of the house. And the problem with that answer would not be that it is not true; it could be a perfectly good explanation by ordinary standards. The problem would be that it is an 'explanation' of how we know some particular fact in the area we are interested only by appeal to knowledge of some other fact *in that same domain*. Hence, that kind of answer could not be *generalized* into a satisfactory answer to the question of how we know anything at all about objects around us, since this way of knowing *already* presupposes knowledge of the subject-matter which was supposed to be under investigation. And if we simply assume from the outset that one has already got some knowledge of the subject-matter under investigation we will not be explaining *all* of it.[36] Hence, it seems that this kind of circularity can be considered as 'harmless' (not vicious) – e.g. by externalist views of justification – only at the cost of being constitutively incapable of satisfying the demand for justificatory generality.[37]

[36] Stroud, *Understanding Human Knowledge*, 101-103.

[37] It has been argued that certain externalist theories of justification can indeed satisfy the demand for justificatory generality without being viciously circular due to their rejection of an epistemological principle that creates the vicious circularity in the first place, namely the internalist view according to which one can know something only by having a reason or argument at his disposal for believing it (William Alston, "Epistemic Circularity," *Philosophy and Phenomenological Research* 47 (1986): 1-30; William Alston, *Epistemic Justification* (Ithaca: Cornell University Press, 1989); Ernest Sosa, "Philosophical Scepticism and Epistemic Circularity," *Proceedings of the Aristotelian Society, Supplementary Volumes* 68 (1994): 263-307). If we reject this internalist principle and hold instead that a belief is justified if it is based on adequate grounds irrespectively of whether the subject knows or justifiably believes this, we can satisfy the demand for justificatory generality with the use of epistemically 'harmless' circular arguments (for example, an argument for the justification of my perceptual beliefs can be epistemically circular in that in forming perceptual beliefs I assume in practice that my perceptual experiences are reliable, but this circularity need not be vicious since I do not have to be justified in making this assumption in order to be justified in the perceptual beliefs that give me my premises). However, I think that this externalist account fails to satisfy the demand for justificatory generality because it cannot discriminate between reliable and unreliable sources of knowledge (William Alston, *Beyond 'Justification'* (Ithaca: Cornell University Press, 2006)), failing thereby to offer a satisfactory understanding of the world and our place in it. It also implausibly implies in this regard that whether one understands how one's knowledge is possible or not depends only on whether the theory he holds about how he came to believe it is true or not. In this way, even if we grant that the externalist account of justification is non-viciously circular, we can do this only at the cost of failing to satisfy justificatory generality, since this account does not enable us to understand that what we have got is reasonable belief in

I argued that justificatory generality can only be obtained in a way that is not viciously circular by satisfying the epistemic priority requirement. I would now want to suggest that the well-known concept of the epistemological *Given* is a *theoretical construct or 'placeholder'* which is created by philosophers (rather than just 'pre-theoretically' 'found' in experience, 'intuition' or 'rational insight') precisely in their attempt to *satisfy* the above epistemological requirements.[38] According to this epistemological picture, there exist certain 'ultimate epistemic atoms' (to which one may or may not have epistemic access) that are justified independently of other 'epistemic units' (i.e. epistemically efficacious conceptual contents), and which, moreover, need not be epistemically supported from evidence provided by the latter in order to be able to transfer their justification to them.[39] The Given, far from being an absolutely transparent, theory free and self-evident concept, is, in essence, a concept laden with philosophical theory which is tailored to satisfy controversial and certainly not self-evident philosophical requirements. Its legitimacy (explanatory power and plausibility) as a concept wholly depends on the legitimacy of the epistemic priority requirement.

However, the epistemic priority requirement can only be 'satisfied' by *scepticism*. The aforementioned theoretical/philosophical constructs which are created for that purpose end up to scepticism, since according to them there exist certain propositional or, in general, representational contents with specific

---

the world's being a certain way (Stroud, *Understanding Human Knowledge*, 146-51)). Notice, for example, that if I am explicitly wondering whether a source is reliable, being told that *if* it is, *then* I have good grounds to believe that it is will be of no help and would certainly not satisfy the requirement of justificatory generality. (For a discussion on the benign nature of certain kinds of circular arguments (those that are rehearsed in a context in which the trustworthiness of a source of knowledge in not put into question) see Michael Bergmann, "Epistemic Circularity: Malignant and Benign," *Philosophy and Phenomenological Research* 69, 3 (2004): 708-26. Yet, the epistemic contexts in which circular arguments are benign do not satisfy justificatory generality nor do Bergmann's arguments about the non-malignant nature of circular arguments in certain contexts purport to establish any connection between benign epistemic circularity and justificatory generality.)

[38] For a thorough examination of the concept of the Given see Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (Cambridge Mass: Harvard University Press, 1997).

[39] At first sight, one might think that only traditional foundationalist concepts (such as the empiricist's 'sense data') can function as 'ultimate epistemic atoms.' But, as a matter of fact, even conceptual systems which are structured according to *coherentist* principles (or, more precisely, the methodological principles which specify the meaning of the concept of coherence in those systems) can perform that function. That is to say, even whole systems of beliefs which at first sight seem to be justified in a thoroughly *holistic* manner, can, in fact, function as 'ultimate epistemic atoms' which are justified independently of the coherence relations in which they may stand to the particular empirical beliefs of the system.

epistemic properties (which can be normative or non-normative in nature and to which the subject may or may not have epistemic access) that are allegedly individuated completely independently of the material inferential relations in which they stand to other epistemic contents (i.e. epistemic contents that are individuated on the basis of their functional role they play in our epistemic practices). Now, if it is assumed that a level of epistemic appraisal can exist which radically *transcends* (i.e. which does not in the least presuppose) that constituted – 'immanently,' so to speak – from the inside of our epistemic practices, it follows that we can never know whether and to what extent our epistemic practices conform to the requirements of the former, inasmuch as a certain *logical gap* necessarily occurs between the *epistemic properties* of the two epistemic levels (level of appraisal and level under appraisal), as a consequence of the fact that the epistemic properties of the former are conceptually *absolutely independent* of the epistemic properties of the latter.

But does this not imply that in the last analysis scepticism is indeed the correct answer to the problem of empirical justification? That would be so if scepticism could indeed count as a genuine and legitimate answer to the problem of empirical justification. Yet, the very definition of the 'atomistically individuated epistemic contents' that was given above is such (radically transcending the epistemic contents which are holistically individuated on the basis of their functional – material-inferential – role in the 'game of giving and asking of reasons') that those 'practice-transcendent' epistemic concepts become explanatorily *idle* (or, to use a Wittgensteinian turn of phrase, they become 'an idle wheel in the mechanism:'[40] any alteration of these 'transcendent' epistemic properties could have no impact whatsoever in the structure of the epistemic contents that are holistically individuated on the basis of the functional role they play in our epistemic practices. Whether the former were constantly changing or did not exist at all, the latter could remain exactly the same. However, in this way *scepticism itself* becomes absolutely idle as a (negative) thesis or stance about the justification of our epistemic practices as a whole, for it is premised – if it is to be regarded as an *epistemic* stance at all – on the possibility of the existence and explanatory relevance of the aforementioned *transcendent epistemic viewpoint*. Only if such a viewpoint genuinely exists and is epistemically connected to our practices can scepticism be considered as a view about the *justification* of the latter, for only then can it *assess them negatively* (i.e. consider them as unjustified) in a meaningful way. Consequently, scepticism cannot really be an *epistemic* view or stance at all; it cannot be the expression of a genuine (negative) judgement

[40] Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1958), §270-71.

about our epistemic position within reality, for this could only be possible if the above transcendent epistemic viewpoint could really function as a premise in an argument whose conclusion would be that we do not have good reasons to think that our epistemic practices (our conceptual systems of empirical beliefs) are justified.

However, it must be noted that there are certain forms of scepticism – e.g. Pyrrhonian scepticism – which do not make claims in support of the viability or truth of their views; they use arguments not to express or justify a certain (sceptical) philosophical viewpoint but in order to expose the problematic premises of the 'dogmatic' philosophical view they wish to undermine.[41] It seems therefore that our above argument against scepticism cannot be applied to this form of resolutely 'non-dogmatic' scepticism since the latter does not put forward any specific or 'positive' epistemological view (not even the view that we cannot know if we ever have good reasons for our empirical beliefs) but is instead functioning 'parasitically' upon certain epistemological views about knowledge and justification, trying to expose the problematic logical consequences of the basic premises those theories themselves (explicitly or implicitly) espouse. Yet, I think that, at the end of the day, even this form of resolutely 'non-dogmatic' scepticism, which withholds judgement about all other stances, cannot escape criticism and is ultimately problematic. Of course, this criticism will not be levelled against the sceptic's epistemological view (for s/he does not have any) but it can be levelled against the sceptic's general epistemological *stance*, which, by staunchly denying to commit to the truth (or the probable truth) of the 'seemings' or 'impressions' that constitute it, ends up being an autobiographical process of recording our subjective impressions, which takes place at a radically non-normative (e.g. causal) level, and therefore cannot be epistemically or, more generally, normatively assessed at all. But if this is so, how can the supposedly 'resolute' 'anti-dogmatic' stance of the Pyrrhonian sceptic be understood (i.e. *assessed*) *as* anti-dogmatic? Is not being 'anti-dogmatic' a *normative* concept? The sceptical stance of withholding commitment about the truth or probable truth of the subject-matter of the sceptic's beliefs (or his opponents' beliefs) seems to entail a rejection of the very 'game of giving and asking for reasons,' where we commit ourselves to the truth of our beliefs and are in turn be held epistemically accountable if our beliefs is not supported by good reasons. This sceptical stance of withholding endorsement seems, at the end of the day, to turn the sceptic into a

[41] Michael Frede, "The Sceptic's Two Kinds of Assent and the Question of the Possibility of Knowledge," in *The Original Sceptics: A Controversy*, eds. Michael Burnyeat and Michael Frede (London: Hackett, 1998), 127-51.

completely 'passive' epistemic subject who refuses to participate in the game of 'giving and asking for reasons.' Yet, this does not make the sceptic resolutely 'anti-dogmatic' but in fact undermines the conditions under which alone can one be considered as 'anti-dogmatic' or as someone who 'withholds judgment' in the first place.

Hence, if, as we argued above, the requirements of justificatory generality and epistemic priority do indeed lead to scepticism, then so much the worse for the former since scepticism does not seem to be an unproblematic thesis or stance whether in modern or ancient guise.

We shall draw this paper to a close by attempting to show the way in which the above sketched explanatory schema can be applied to BonJour's, Haack's and Dancy's theories of justification and explain, at a deeper level, their inevitable inadequacies.

In BonJour's case, things are relatively straightforward: he explicitly accepts the challenge posed by the epistemic priority requirement and attempts to provide a straight solution to it, not only in his 'late' foundationalist period (as one might expect), but also in his 'early' coherentist phase. More specifically, the acceptance of the epistemic priority requirement leads BonJourean cohrerentism to the view that the question of metajustification – i.e. of the justification of a whole system of empirical beliefs – is perfectly legitimate, and accordingly creates a certain theoretical pressure towards the view that the an effective response to the question of metajustification could only be formulated from an epistemic standpoint outside coherence altogether. This is also why he thinks that only a robust correspondence theory of truth can provide a successful answer to the question about the nature and function of truth.[42] Again, it is for this same reason that he considers one of his key epistemological concepts, that of the doxastic presumption, to be a failure; it fails to vindicate coherentism about empirical justification because, ultimately, it cannot be itself justified independently of coherence considerations. Indeed, BonJour's radical change of epistemological perspective, namely his conversion from coherentism to foundationalism, is readily explained, as he himself suggests, by his persistent efforts and ultimate failure to justify the doxastic presumption independently of coherence considerations. And, as BonJour correctly recognises, the only way to do this without violating the epistemic priority requirement is to opt for a purely foundationalist justification of the doxastic presumption.[43]

[42] BonJour, *Coherence Theory*, 171.

[43] See e.g. BonJour, "The Dialectic of Foundationalism and Coherentism"; BonJour, "Foundationalism and the External World," 229-49.

Now, in contrast to BonJour (both 'early' and 'late'), Dancy's aim is to develop a (coherentist) view of justification which explicitly *rejects* the epistemological requirements of justificatory generality and epistemic priority. However, although Dancy indeed seems to resolutely reject those 'twin' problematic epistemological requirements in his philosophical practice, in my view, at the end of the day, he does not succeed in being completely immunized from their subterranean theoretical influence. For example, a strong indication of the subliminal influence that the epistemic priority requirement exerts over Dancy's coherentism (despite his avowed intentions to the contrary) is that he takes it that a necessary corollary of his pure coherentism about justification is a notion of empirical truth which, on the one hand, is supposed to be defined in terms of (explanatory) coherence – i.e. in terms of an 'ideally coherent system of beliefs' – (hence, no sceptical epistemic gap seems to have been opened between empirical justification and truth), while, on the other hand, it is also defined in a way that leaves open the possibility that the ideally coherent system, being essentially something irreducible to our epistemic 'finitude,' can transcend every possible (essentially 'finite') coherent system which purports to represent reality; worse, given the definition of 'ideally coherent system' it is always possible that an increase in the system's explanatory coherence will not be accompanied with an increase of the system's reliability in representing external reality. It is my contention that this eventual unexpected sceptical turn of Dancy's coherentism is in large part the result of his unintended adherence to the view that only something radically independent of coherence relations among beliefs could function as their truth-maker. That is, he seems to be pressed in certain central points of his own version of 'metajustification,' and despite his protestations to the contrary, to accept the legitimacy of the epistemic priority requirement. This is why although he officially has to hold that justification (explanatory coherence) is internally related to truth, at the end of the day, when he explicates what this amounts to, this internal relation can hardly be distinguished from an external relation between justification and truth, of the form that classic correspondence theories of truth.

Haack's 'hybrid' foundherentist theory of justification, exactly like Dancy's coherentism, does not explicitly accept the justificatory generality requirement (and its corollary – the epistemic priority requirement), since her preferred solution to the problem of the infinite regress of justification essentially involves *coherentist* elements (which are not fitted to satisfy the epistemic priority requirement). Hence, in contrast to BonJour, Haack, like Dancy, does not seem to think that the epistemological requirements in question are legitimate and in need

of an urgent answer. She would probably argue that her conciliatory epistemological position provides answers which conceptually transform the traditional epistemological puzzles (infinite regress of justification, metajustification) that appear as meaningful and urgent questions precisely by rejecting the epistemological requirements of justificatory generality and epistemic priority. Now, I would not want to dispute Haack's willingness to move towards this direction, neither do I want to downplay the important insights offered by her foundherentist theory as the latter progressively unfolds in her seminal book *Evidence and Inquiry*. But I think that she is not as resolute in her rejection of those twin problematic epistemological requirements as she should be. For example, Haack attempts to solve the problem of securing that our conceptual system of beliefs actually receives input from extra-conceptual reality with the use of essentially foundationalist or 'Givenist' conceptual tools (namely, the initial positive epistemic status which non-conceptual S-evidence possess independently of their justificatory support from conceptually structured reasons) without attempting at the same time to change the meaning of those foundationalist epistemic features of her theory so as to extinguish their 'Givenist' connotations. I take it that this just shows that she has not in fact overcome or transcend the conceptual frame in which the justificatory generality requirement is formulated, gets its meaning and (accordingly) demands an urgent and direct answer (or else a withdrawal to scepticism). Despite the fact that Haack would in all probability not want to consider this requirement as legitimate and in need of an urgent answer, her own philosophical practice does not successfully support this claim at least to the extent to which she does not radically transform – as she must if she really had exorcised the epistemic priority requirement – the very meaning of foundationalist conceptual tools on the basis of which she attempts to solve the problem of securing that our conceptual system actually receives input from extra-conceptual reality. And it is in large part for this reason that Haack's solution to this traditional epistemological problem cannot be considered as fully successful.[44]

[44] See e.g. the critique of Haack's foundherentist theory in section 4 where I argue that the foundationalist inspired 'non-doxastic' sources of justification (sensory S-evidence), which Haack thinks are necessary for a proper solution to the problem of external input, in fact only provide a verbal 'solution' to this problem. Those non-doxastic epistemic sources end up being epistemically idle in the overall context of Haack's theory; they cannot transmit their 'partial' (but supposedly independent) justification to the conceptual-doxastic level (i.e. to our belief-system) without themselves receiving justificatory support from (doxastic) epistemic sources of the latter. This is because on the one hand, Haack wants to preserve a radical independence of the 'evidence of the senses' from the conceptual level of our system of empirical beliefs (this is why she thinks it is necessary to interpret them as resolutely non-doxastic, non-conceptual

7. Concluding Remarks

How can we properly conclude our examination of the thorny philosophical issue of empirical justification after all the above critical remarks to BonJour's, Haack's and Dancy's theory? Do the problematic epistemological assumptions that as we argued are intimately involved in all four theories that were examined in this paper lead to any positive results or proposals about empirical justification or are they only useful for the provision of a thoroughly 'negative' critique of the problem in question? Within the confines of the present paper, whose arguments were of a purely negative character, such a positive proposal obviously cannot be adequately developed. However, it can be stressed that the unacknowledged background epistemological assumptions that were unearthed in the theories of justification under examination (i.e. the requirements of justificatory generality and epistemic priority as well as the concept of the epistemological Given) and the arguments that were used against them do indeed point towards the formulation of a positive proposal about empirical justification. An example of this is what might be called 'contextualist functionalism,' according to which the justificatory status of one's empirical belief (e.g. whether this belief can be epistemically efficacious without being inferred from other beliefs or only by being inferred from the latter) depends on the *functional role* of this belief within an essentially socially structured *logical space of reasons*. This view completely rejects the epistemological requirements of justificatory generality and epistemic priority as well as the concept of the epistemological Given in all its forms,[45] and thereby

epistemic sources), while on the other hand, she also believes that the justificatory contribution of doxastic-conceptual and holistic epistemic sources to the non-doxastic level is necessary in order for the latter to be able to transmit whatever justification it possesses to the former. She makes this latter move (which, at first sight, seems to imply a *rejection* of the epistemic priority requirement) in order to avoid a traditional foundationalism, but in the end, she does not succeed in solving the problem at hand since, by maintaining that experiential S-evidence must have a positive epistemic status independent of all conceptually structured reasons, she seems to want to *preserve* or at least not wholeheartedly reject the epistemic priority requirement, which, in turn, as we saw, can be satisfied only by pure foundationalist positions. That is to say, she does not seem to really change the meaning of the foundationalist and coherentist elements that are preserved in her theory in a way that would avoid the above radically ambivalent stance towards the need to satisfy the epistemic priority requirement in a successful theory of empirical justification.

[45] Note that the rejection of the requirements of justificatory generality and epistemic priority is not equivalent to the rejection of the (correct) view that an explanans can properly account for the explanandum only if its content is independent of the latter. Consider the case of scientific causal explanations: although their content is (and ought to be) independent of the phenomena to be explained, this does not mean that the former are thereby understood as independent of

differs from foundationalism, (pure or impure) coherentism and foundherentism alike.[46] However, the proper development and defence of this alternative view about empirical justification will have to wait for another paper.

---

the latter in a foundationalist sense of the term (according to which a reason for believing something can be good, non question-begging only if its content satisfies the requirements of justificatory generality and epistemic priority). For example, it can be argued that scientific causal explanations can function as reasons (for placing the phenomena they explain in 'the logical space of causes') due to their contribution in the diachronic process of the *self-correction* of our system of empirical beliefs (i.e. a correction based on potentially revisable standards that are internal to the practices in question) -and not because they are 'independent' in an absolutely presuppositionless sense of the term (see e.g. Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (Cambridge Mass: Harvard University Press, 1997), §38).

[46] Notice, for example, that unlike (pure or impure) coherentism and foundationalism, this alternative view does not construe the epistemic *dependence* of a belief on certain conditions in terms of this belief being *inferable* from the presence of those conditions – or being *coherent* with beliefs about the conditions in question (see also Sellars, *Empiricism and the Philosophy of Mind*, §32). More specifically, there is a distinction between a background of epistemic practices being required for the possibility of my knowing e.g. that something is red and my inferring that it is red from a belief about those background practices. Any piece of knowledge may depend on a whole host of background conditions such as an agent's perceptual or linguistic skills or her implicit knowledge of generalities but it does not follow that this kind of dependence is to be understood as inferential in nature (see also Matthew Burstein, "Prodigal Epistemology: Coherence, Holism and the Sellarsian Tradition," in *The Self-Correcting Enterprise: Essays on Wilfrid Sellars*, ed. Michael P. Wolf and Mark Norris Lance (New York: Rodopi, 2006), 202-207). Indeed, if the above background conditions are satisfied one can have direct, non-inferential justification of, say, the presence of certain objects and events in his surroundings, while, at the same time, this justification can well be dependent on our motor, perceptual or linguistic skills, and most importantly, on the application of a certain (non-sacrosanct) categorial framework (or theoretical paradigm) in the world.

DEFENDING STANDARDS CONTEXTUALISM

Robert HUDSON

ABSTRACT: It has become more common recently for epistemologists to advocate the pragmatic encroachment on knowledge, the claim that the appropriateness of knowledge ascriptions is dependent on the relevant practical circumstances. Advocacy of practicalism in epistemology has come at the expense of contextualism, the view that knowledge ascriptions are independent of pragmatic factors and depend alternatively on distinctively epistemological, semantic factors with the result that knowledge ascriptions express different knowledge properties on different occasions of use. Overall, my goal here is to defend a particular version of contextualism drawn from work by Peter Ludlow, called 'standards contextualism.' My strategy will be to elaborate on this form of contextualism by defending it from various objections raised by the practicalists Jason Stanley, Jeremy Fantl and Matthew McGrath. In showing how standards contextualism can effectively repel these criticisms I hope to establish that standards contextualism is a viable alternative to practicalism.

KEYWORDS: pragmatic encroachment, standards contextualism, practicalism, Jason Stanley, Jeremy Fantl, Matthew McGrath

1. Introduction

According to the proponents of the pragmatic encroachment on knowledge, whether one can be said to know a claim depends on the practical circumstances in which one finds oneself. For example, according to Jason Stanley's 'Interest-Relative Invariantism,' "whether or not someone knows that p may be determined in part by practical facts about the subject's environment."[1] Jeremy Fantl and Matthew McGrath defend a similar view. They defend a pragmatist principle called 'Action' which states: "if you know that p you are proper to act on p when the question of whether p is relevant to the question of what to do."[2] Stephen Grimm calls the view Stanley, Fantl and McGrath are defending 'practicalism,' which Grimm contrasts, "borrowing Stanley's label and basic idea, and in keeping with similar thoughts by Fantl and McGrath," with the doctrine of

[1] Jason Stanley, *Knowledge and Practical Interests* (New York: Oxford University Press, 2005), 85.

[2] Jeremy Fantl and Matthew McGrath, *Knowledge in an Uncertain World* (New York: Oxford University Press, 2009), 49.

'intellectualism,' that "whether a true belief amounts to knowledge depends exclusively on truth-related factors."[3]

The defense of practicalism, Grimm notes, usually focuses on the celebrated 'bank cases.'[4] In general terms the bank cases involve an agent who has seemingly good evidence for a true claim, but who intuitively speaking knows this claim, or not, depending on how much is at stake for her. That is, given the same amount of evidence, it can happen that the agent knows this claim if very little is at stake, but not know the claim if a lot is at stake. For Stanley, Fantl and McGrath, and Grimm, the best way to explain this phenomenon is to reject intellectualism (or more precisely, for Grimm, 'threshold' intellectualism) and to allow into the normative evaluation of knowledge claims practical considerations.

But, as Stanley points out, the celebrated bank cases have also been used to justify contextualism, the "distinctively epistemological"[5] semantic thesis that knowledge predicates (i.e., 'knows that …') "denote different knowledge properties on different occasions of use."[6] What it means for context-sensitivity to be distinctively epistemological is subject to interpretation. "A sentence is context-sensitive," Stanley asserts, "if and only if it expresses different propositions relative to different contexts of use."[7] So, with distinctively epistemological context-sensitivity, we find that epistemological factors lead to sentences expressing different propositions. But how does this come about? Stanley describes various ways this can occur, but for the most part settles on one version (deriving he notes from work by Peter Ludlow[8]). On this version,

> predicates that are instances of the schema 'knows that p' are context-sensitive since they are really of the form 'knows that p relative to standards s' where s receives a value from context.[9]

I will call this version of contextualism 'standards contextualism.'

An important feature of contextualism, generally speaking, is that it is a form of intellectualism: with contextualism, whether one knows a claim, or not, is determined independently of pragmatic factors. Thus, practicalists like Stanley,

---

[3] Stephen Grimm, "On Intellectualism in Epistemology," *Mind* 120 (2011): 706.

[4] Grimm, "On Intellectualism in Epistemology," 707.

[5] Stanley, *Knowledge and Practical Interests*, 16.

[6] Stanley, *Knowledge and Practical Interests*, 3.

[7] Stanley, *Knowledge and Practical Interests*, 16.

[8] Peter Ludlow, "Contextualism and the New Linguistic Turn in Epistemology," in *Contextualism in Philosophy*, eds. G. Preyer and G. Peters (Oxford: Oxford University Press, 2005).

[9] Stanley, *Knowledge and Practical Interests*, 17.

Fantl and McGrath regard contextualism as a foil to their pragmatist conceptions of knowledge, and accordingly each provides arguments against contextualism. My plan in this paper, in elaboration of standards contextualism, is in section 2 to defend it from objections raised by Stanley, and in section 3 to defend it from objections raised by Fantl and McGrath. I then examine and respond to two further objections to standards contextualism in the final section. Overall, I hope to show that standards contextualism is a viable alternative to practicalism.

2. Stanley Versus Contextualism

Stanley provides a critique of contextualism from three vantage-points which we examine in turn. He considers, first, the bank cases and argues that contextualism fails to get the right answer with a certain version of these cases. Second, he introduces various 'linguistic' considerations and argues that, with contextualism, we wrongly conclude that knowledge ascriptions are gradable and that certain anomalous speech-act reports and anaphora are acceptable. Lastly, he takes aim at Ludlow's unique brand of standards contextualism which he believes wrongly assigns a position for standards in all kinds of predications, not just in epistemic ones (i.e., knowledge ascriptions). We examine and respond to each of these criticisms.

2.1 The Ignorant High Stakes Bank Case

In *Knowledge and Practical Interests* Stanley examines five versions of the bank cases. In all these versions, the situation concerns an agent who is deliberating about whether to stand in line at a bank on a Friday to deposit a cheque or wait to deposit the cheque the next day. The question is whether she can be said to know the proposition, "the bank will be open tomorrow (Saturday)." The common view is that whether the agent knows the proposition depends on what's at stake for the agent should her belief be mistaken. In one case, called 'Low Stakes,' there is little at risk for the agent since she has no impending bills due, and as such the evidential facts are sufficient for her to be said to know that the bank will be open on Saturday. By comparison, in the 'High Stakes' case, there is much at risk for the agent if she is mistaken (say, she has a bill coming due and the money needs to be in her account by Monday morning), and so with the same evidential facts she is said not to possess knowledge. According to Stanley, contextualism effectively handles these sorts of cases.[10] But the situation is different for the case called 'Ignorant High Stakes.' In this case someone thinks she is in a low stakes situation

[10] Stanley, *Knowledge and Practical Interests*, 24.

but is actually in a high stakes situation, and so wrongly ascribes to herself knowledge. It is Stanley's opinion that contextualism gives the wrong result in this sort of case because it incorporates the following claim:

> what determines the semantic value of instances of 'knows that *p*', relative to a context of use, is some collection of facts about the intentions and beliefs of the conversational participants in that context of use.[11]

Accordingly, because in Ignorant High Stakes the agent (wrongly) believes that the stakes are low, the agent expresses a knowledge claim using epistemic standards that are themselves low. Thus, if the semantic value of this knowledge claim is set by these low standards (presumably as per standards contextualism) the agent can be said to possess knowledge.

Motivated by the intuition that the agent lacks knowledge in Ignorant High Stakes, Stanley proposes a theory of knowledge in which the meaning of 'knows that p' does not vary with the context ('invariantism') and that whether an agent knows that p varies with whether p is a serious practical question for the agent (knowledge is 'interest-relative'). A "serious practical question," for Stanley, is a "proposition that one must take into account in decision making," where one's obligation to take a proposition into account varies with how the truth of this proposition affects the warranted expected utility of the actions at one's disposal.[12] As he further explains, one makes use of warranted expected utilities, and not alternatively 'subjective credences,' "because the agent might not be aware of what is in her own best interest."[13] That is, the warranted expected utility of an action is an objective quantity, but may not even be objective enough for Stanley. He comments:

> warranted expected utility is probably not sufficiently impersonal of a notion to do the required work. There may be facts relevant to the utility calculation that the agent is not epistemically responsible for knowing. So a more impersonal notion of utility may be required to capture the notion of a serious practical question.[14]

The upshot is that an agent may be unaware that a proposition is a serious practical question for her, and so is mistaken in thinking she knows a claim. This is the sort of situation Stanley asserts we have in Ignorant High Stakes.

A standards contextualist can respond to this problem in the following way. She can suggest that the agent is in a position to know the relevant claim in

[11] Stanley, *Knowledge and Practical Interests*, 23.
[12] Stanley, *Knowledge and Practical Interests*, 94.
[13] Stanley, *Knowledge and Practical Interests*, 95.
[14] Stanley, *Knowledge and Practical Interests*, 95.

Ignorant High Stakes, given her low (but not unreasonable) standards. Accordingly, if she were to bring about a belief in the claim, and assuming the claim is true, she could be said to know the claim. However, given the practical situation, the standards contextualist can further suggest that, practically speaking (though not necessarily epistemically speaking), the agent should be using higher standards if there is a severe cost should the claim be false. Moreover, with the adoption of these higher standards, it may be that the agent isn't in a position to know the claim, after all. The standards contextualist may then recommend that the agent not bring the belief about, given the costs in being wrong, with the result that she doesn't know the claim (even) in the low stakes situation since she doesn't believe the claim.

This is in fact the sort of approach to Ignorant High Stakes that I suggest we should take, and which I develop more fully later on. It is an approach that asks us to distinguish between the decision that, under certain evidential circumstances, we should believe a proposition (though not necessarily bring the belief about), and the decision that, where further practical considerations are taken into account, we should go further and actually bring about a belief in this proposition. In order to motivate this approach, let us look at some potential concerns with Stanley's practicalist approach to Ignorant High Stakes.

To begin with, what makes a proposition a serious practical question for Stanley is an objective matter, one that is determined independently of what an agent is even 'epistemically responsible for knowing.' But it isn't, nor could it be, exclusively an objective matter, a point Stanley seems to acknowledge. Consider the proposition that you have an odd number of hairs. Stanley says,

> given that I do not care about the number of hairs you have, whether or not you have an odd number of hairs will not make a difference to the warranted expected utilities of retaining or discarding my belief. So, the proposition that you have an odd number of hairs is not a serious practical question for me.[15]

Going back to Ignorant High Stakes, then, it might well be the case that the agent doesn't care about the practical matter at hand – specifically, as the case is described, about whether a certain impending bill is coming due.[16] As such, because the agent doesn't care, she will possess knowledge after all for Stanley since it will become a low stakes situation for her. Of course, the scope of things people care about is changeable, and the agent may change her mind periodically about whether pending bills are a concern. It follows that whether the agent knows, or not, fluctuates with whether she cares about impending bills, and these

[15] Stanley, *Knowledge and Practical Interests*, 96.

[16] Stanley, *Knowledge and Practical Interests*, 5.

two matters seem totally unrelated to one another. It may even seem 'mad', as Fantl and McGrath – admitted practicalists – confess: "what is mad is the idea that whether you are in a position to know could be affected by stakes. But that is precisely what the pragmatist approach requires."[17] If one wants to gain knowledge, it seems one need only strive to not care about the truth of the claim one is considering.

In order to restore some order as regards what one knows, to make what one knows less dependent on one's changeable set of values, one may decide to seek some normative standards that regulate a prospective knower's values. For instance, in Ignorant High Stakes, it may be that the agent is foolish in not caring about impending bills: she should care about them given how they may impact her life. Comparatively, the number of hairs someone has is obviously, and objectively irrelevant to one's practical concerns (assuming a normal circumstance), and it may be that this is Stanley's point in introducing this case. But the issue of what values one should have is not a straightforward, meta-ethical matter. There is plenty of debate about what things deserve to be valued, and this is certainly a matter that epistemologists should not feel the obligation to express an opinion about. The situation is further complicated by the fact that people often have values that run in different directions. In Ignorant High Stakes, the agent may care about impending bills, and so lack knowledge from that perspective, while also greatly value the appearance of being someone who knows things, and so from that perspective possesses knowledge. What then should a practicalist say about a situation where a knowledge ascription has both positive and negative practical implications? One can imagine many similar cases where an agent's values pull the agent in opposite directions, where there is ambiguity (from a practicalist perspective) on the question of whether an agent knows. Indeed, the ambiguity here becomes more complicated once one considers that an agent may have even more values that pull in yet other directions. One lesson here is that we should be cautious about using cases that are over-simplified. In the original bank cases, the matter turns solely on whether the agent has impending bills due or not, and if so whether she should put off or not going to the bank on Saturday. But practical decisions can get very complicated – a person's value set can contain many diverse elements. The important point is that these valuational complications can have crucial epistemic implications for a practicalist like Stanley, no matter what they are and no matter how trivial they seem, and that this seems to take us beyond what should be the proper ambit of epistemology.

[17] Fantl and McGrath, *Knowledge in an Uncertain World*, 27.

To further illustrate how unnecessarily complicated matters can become, consider once more High Stakes, the case in which there is much at stake if the agent is mistaken. The case seems straightforward for the practicalist: there's lots at stake so the agent has to be sure her belief is correct, and so despite the fact that with the very same evidence she can be said to know in the low stakes case, she may not know when the stakes are high. But the same thing may happen in this case as we find in Ignorant High Stakes – the agent may be mistaken about the stakes and so we have a parallel Ignorant Low Stakes case in which the stakes are low, though the agent thinks they are high. What this means is that in any particular situation, since an agent may be mistaken about the stakes, she might be mistaken as well about whether she knows. So with practicalism we have the somewhat puzzling predicament that, despite the constancy of the evidence, it remains unclear whether an agent knows a claim if there is uncertainty about the practical situation. An agent knows or doesn't know, unbeknownst to her, dependent on what the hidden stakes are, without any changes in the evidence, in the truth of the claim, or in the agent's state of belief. To paraphrase Stanley, the practicalist can handle this situation "only at the cost of advancing a rather dramatic claim about the potential [epistemic] effects of non-psychological facts about extralinguistic [stakes]."[18]

So our general conclusion about the Ignorant High Stakes case is this: whereas Stanley maintains that Ignorant High Stakes is a case the contextualist has trouble with, we have argued that the case poses just as much, if not more trouble for Interest-Relative Invariantism. So having then defused this potential, practical problem for (standards) contextualism, let us now look at the linguistic critiques Stanley offers against contextualism.

2.2 Stanley's First Linguistic Critique: Knowledge Ascriptions are Not Gradable

Stanley notes that for contextualists such as Stewart Cohen and Keith DeRose, "knowledge ascriptions come in varying degrees of strength," that is, "they are intuitively gradable."[19] This is what we would expect from standards contextualism where the standards that modify knowledge ascriptions could be high (more demanding) or low (less demanding). As such, if (standards) contextualism were true, we'd expect linguistic expressions of knowledge to pass two tests: (1) they should allow for modifiers, and (2) be amenable to comparative constructions. Stanley asserts that since knowledge ascriptions fail both tests, they

[18] Stanley, *Knowledge and Practical Interests*, 26.
[19] Stanley, *Knowledge and Practical Interests,* 35.

are not gradable after all, and so (standards) contextualism is false.[20] In looking at these tests in more detail I plan to show that, properly understood, knowledge ascriptions pass these tests.

Consider first the use of modifiers. Stanley uses as paradigm examples of acceptable modifications of gradable expressions "Michigan is flat, but not really flat"[21] and "'I don't like Bill very much."[22] Can the word 'know' be used in analogous expressions? Stanley thinks not. Consider the sentence, "John knows that the bank is open, but doesn't really know that the bank is open," which Stanley finds extremely odd. Similarly, consider "I don't know very much that Bush is president," which also sounds peculiar. A possible diagnosis for why these expressions sound odd is that knowing mundane facts, such as whether the bank is open or whether Bush is president, is a 'yes/no' sort of issue – either you know it or you don't –, and in fact all of Stanley's examples are of this kind. But there are areas of inquiry where states of knowledge are more nuanced, such as in the sciences. For example, a student might say, "I know that atoms have orbitals (say, well enough to pass the test), but I don't really know that atoms have orbitals." What the student is saying is that she is aware of the basic fact of orbitals, but cannot provide the fine details of orbital theory, which would require a deeper understanding of atomic physics. That is, on a lower standard (the standard used in assessing students), she knows that atoms have orbitals, but on a higher standard (the standard used by professional physicists), she doesn't (really) know this claim. Here is a less technical example: someone with a stuffy nose and a sore throat says, "I know I have a cold, but I don't really know I have a cold." For the purposes of day-to-day discourse, telling people that one has a cold is sufficient (say, to distinguish one's ailment from allergies), but not adequate if one were to seek medical precision. For example, although a cold is similar to the flu it is in fact much different, and so to know on a more rigorous standard that one has a cold one needs to recognize how it differs from the flu. Again, one may have actually gotten rid of a cold, yet the symptoms may be lingering, leading one to think that one still has a cold. Because of these complications, one may know one has a cold on a looser evidential standard, but not *really* know one has a cold if one adopts a more rigorous standard. Nevertheless, loose standards as regards one's state of knowledge may be completely appropriate, such as when one enters the office in the morning and is making light conversation. Sniffling and sneezing, one

[20] Stanley, *Knowledge and Practical Interests*, 36-45.
[21] Stanley, *Knowledge and Practical Interests*, 37.
[22] Stanley, *Knowledge and Practical Interests*, 38.

is asked whether the source might be allergies, to which, "I know I have a cold. It's not just allergy, so you'd better keep your distance," is a perfectly reasonable comment, despite one's ignorance of the differences between a cold and the flu. But then if pressed about the difference between a cold and the flu, one might confess, "I know I have a cold, but (ok) I don't *really* know I have a cold, now that you insist on the difference between a cold and the flu."

"I don't know very much that Bush is president" is an unusual sentence, but that may be because of the current state of English and not a reflection of the non-gradability of knowledge claims. More sensible sounding is, "About the claim that Bush is president, I don't know very much." Still, it is an awkward comment because there is not much to know about the claim that Bush is president: either he is president or not, and either one knows this or not. But take instead a more nuanced claim, such as the one we used above: "About the claim that atoms have orbitals, I don't know very much." This claim would be true if one's knowledge that atoms have orbitals is simplistic and meets the lowest of standards, such as if one just believes this on the basis of one's faint recollection of a high school chemistry book. On the other hand, if one is a chemistry professor, one would need to know well that atoms have orbitals. A chemistry professor would very much know that atoms have orbitals in that she understands orbital theory, and particularly understands the evidential basis to this theory. Indeed, an indicator of how well one knows this claim could be one's ability to defend this claim from critique. Similarly, one would very much know that one had a cold if one understood its viral nature and thus its insensitivity to antibiotics. One would know less well that one had a cold if one thought one could be cured by taking antibiotics.

One may nevertheless resist these arguments on the basis of the awkwardness of the resultant expressions. It may seem that I am trying to force the English language to comply with the dictates of standards contextualism. Of course, all languages, including English, are changeable, and new grammatical, stylistic constructions are now the norm with the fluidity and expansion of technologically-enhanced means of communication. What were once awkward expressions can subsequently become highly acceptable. This has, in fact, already happened in epistemology. For example, Bayesians talks about 'degrees of belief,' where one can strongly believe a claim (i.e., assign it a high probability) or weakly believe it (i.e., assign it a low probability). Thus, one would say, "I believe that I have a cold with a probability of .9," which is about as awkward a sentence as one would ever find. But that infelicity has not hampered Bayesianism as a viable epistemology, and for some it is even a favoured, normative theory of belief.

What about the use of 'know' in comparative constructions? A comparative construction involving the gradable verb 'like' that Stanley regards as acceptable is, "John likes Bill more than Mary does." An analogous construction with the verb 'know' does not sound at all acceptable: "John knows that Bush is president better than Mary does."[23] Stanley then argues, "if the semantics of 'know' did involve scales of epistemic strength, then there should be uncontroversial examples of non-idiomatic comparison and modifications,"[24] and since there are not, knowledge ascriptions are not gradable. But just as we can find acceptable expressions using modified uses of the verb 'know,' so we can find some uncontroversial comparative constructions involving 'know.' Here's one: "Professor X knows that atoms have orbitals better than her students do." Again, the example involving Bush sounds odd because knowing that Bush is president is an uncomplicated 'yes/no' matter. On the other hand, more sophisticated claims can be known with more or less intellectual rigour. It sounds reasonable to say that Professor X's knowledge of atomic orbitals is better than her students' knowledge in that she has an awareness of the evidential basis to this claim and how this claim fits into the overall explanatory structure of atomic theory. Students, by comparison, would simply know this claim by rote. Note that the sense in which the professor knows better than atoms have orbitals is not simply that she possesses more facts about atomic orbitals. It may be that John knows more facts about Bush and about his presidency than Mary does, but still does not know better than Mary that Bush is president. The difference with the question of Bush's presidency is that the evidential basis to knowing that Bush is president is straightforward and uncomplicated, not requiring the use of sophisticated, experimental apparatus. Moreover, possessing this knowledge doesn't require a lot of theoretical complexity: simply, the presidential office is the highest executive office in the land, and Bush occupies the post. One knows just as well that Bush is president regardless of one's comprehension of the details of the American political system or of how one acquires this knowledge through media sources. To take another mundane claim, one knows that something is a car regardless of one's understanding of the internal combustion engine, the makes and models of cars built by major automobile companies, and so on. Thus, it sounds awkward to say that someone knows better than another that something is a car. By comparison, one could know better that one has a cold than another person if one is aware of the subtle symptomatic differences between a cold and the flu that allows one to distinguish them. So, to summarize what's been argued in this section, with

[23] Stanley, *Knowledge and Practical Interests*, 40.
[24] Stanley, *Knowledge and Practical Interests*, 40.

(standards) contextualism we'd expect the word 'know' to be gradable in that one can know a claim to a greater or lesser degree. As such, we'd expect linguistic expressions of knowledge to allow for modifiers, and be amenable to comparative constructions. Although we concede that knowledge expressions of the more mundane, unsophisticated sort typically do fail gradability (in not allowing for modifiers and not being amenable to comparative constructions), this is not the case with more complex assertions of knowledge, such as those found in the sciences. With these more complex matters, one can know claims to a greater or lesser degree, depending on one's awareness of the evidential basis to these claims and one's understanding of the explanatory theories that underlie them. As the sciences are commonly held to be our prime repository for first-class knowledge claims, the gradability of scientific knowledge claims speaks on behalf of standards contextualism.

2.3 Stanley's Second Linguistic Critique: Unusual Speech-Act Reports and Anaphora

Stanley's second critique of contextualism involves some sample conversations people would have if contextualism were true, conversations Stanley finds to be highly problematic. In the first conversation, A and B are looking at a zebra in a zoo and A asserts, "I know that is a zebra." B points out that the animal is indistinguishable (for A) from a cleverly painted mule, which A concedes. B then second-guesses A's pronouncement that she knows it is a zebra, to which A responds, "I didn't say I [knew it was a zebra]."[25] Stanley finds this speech act report to be "very strange" (indeed, "well-nigh incoherent"[26]), though it sounds "perfectly reasonable" to the contextualist.[27] In the second example, someone is reflecting on whether she knows that she as hands. "If I have hands, then I know I have hands," she says to herself. She then considers the skeptical possibility that she is a brain in a vat, and in the midst of seriously considering this possibility comes to the conclusion that, even if she has hands, she doesn't know that she does. Nevertheless, she concedes, "what I said earlier is still true," where 'what I said earlier' anaphorically connects to "if I have hands, then I know I have hands." Stanley finds such an anaphor "very difficult to grasp,"[28] even though a contextualist would find it unproblematic.

---

[25] Stanley, *Knowledge and Practical Interests*, 52.
[26] Stanley, *Knowledge and Practical Interests*, 56.
[27] Stanley, *Knowledge and Practical Interests*, 52.
[28] Stanley, *Knowledge and Practical Interests*, 54.

Stanley is frank that these sorts of discourses are perfectly sensible with other sorts of contextually sensitive terminology, such as with 'possibility'[29] and 'wealth.'[30] What is not made clear by him is why the word 'know' is ineligible for such context-sensitivity. My suspicion is that, whereas the examples Stanley provides illustrating the context-sensitivity of 'possibility' and 'wealth' are familiar to everyday speakers, the examples relating to the word 'know' describe skeptical possibilities that many people find fanciful, if not ludicrous. The possibilities that zoos are populated with zebra pens filled with cleverly painted mules (as though mules look at all like zebras), or that we might be brains in vats, are not terribly serious. No one is going to doubt their knowledge that they're seeing zebras or that they have hands solely on such bases. So to test Stanley's intuitions about the failure of context-sensitivity as regards knowledge claims, we need to find an example that does not trade in extraordinary skeptical scenarios. Here is one such case.

As is well known, crocodile and alligators are quite similar. In fact, most people do not know how to tell them apart. For those of us who don't live in areas where these animals are endemic, or who are not biologists, the words are likely interchangeable: a 'crocodile' is a 'crocodile or alligator.' Now suppose A and B are again at the zoo, this time near the crocodile pool, and A asserts about a crocodile, "I know that is a crocodile." B points out that the animal is indistinguishable (for A) from an alligator, which A again concedes. B then second-guesses A's pronouncement that she knows it is a crocodile, to which A responds, "I didn't say I knew it was a crocodile." Is A's speech act report 'very strange' and 'well-nigh incoherent'? I would say it is perfectly reasonable. In essence, A's initial pronouncement is uttered on the basis of the low standards appropriate to those for whom crocodiles are indistinguishable from alligators (but quite distinguishable from snakes, frogs and so on). B is then pointing out that A's categorization is too loose, and that because she doesn't know the difference between crocodiles and alligators, she doesn't really know – using higher, more scientific standards – that the animal she is looking at is a crocodile. A's speech act report, then, amounts to the admission that she wasn't using these higher standards in saying that the animal is a crocodile.

The anaphoric case is dealt with similarly. Suppose someone is reflecting on whether she knows that an animal is a crocodile. She thinks, "If that is a crocodile, then I know it is a crocodile." She then considers the non-skeptical possibility that the animal is actually an alligator, and in the midst of seriously considering this

[29] Stanley, *Knowledge and Practical Interests*, 53.
[30] Stanley, *Knowledge and Practical Interests*, 56.

possibility comes to the conclusion that, even if it is a crocodile, she doesn't know that it is. Nevertheless, she concedes, "what I said earlier is still true," where 'what I said earlier' anaphorically connects to "if that is a crocodile, then I know it is a crocodile." Is such an anaphor 'very difficult to grasp'? Not at all, once we see her earlier pronouncement as uttered in the context of the lower standards people sometimes adopt as regards what counts as a crocodile, where on such lower standards crocodiles and alligators are essentially the same kind of animals. In effect, she recognizes that on a more rigorous standard her claim to know that the animal is a crocodile is disputable, but affirms nevertheless that she could still be said to know on a common, looser standard.

So far we have examined, and responded to Stanley's objections to contextualism on the basis of contextualism's alleged failure to generate the right result in Ignorant High Stakes, as well as its tendency to produce awkward linguistic constructions. We now look at one further criticism of contextualism, specifically, Stanley's misgivings about the sort of contextualism we are advocating here, 'standards' contextualism. Stanley's focus is standards contextualism as advanced by Peter Ludlow,[31] to which we now turn.

2.4 Ludlow's Standards Contextualism

As Stanley recounts, Ludlow highlights the fact that references to standards are common in scientific discourse. To illustrate, Stanley cites the following examples:

1) John doesn't know that water is a liquid by the standards of chemistry.

2) Copernicus didn't know that the sun was at the centre of the solar system by today's standards of knowledge.[32]

Now it's worth pointing out that these examples are somewhat unclear. First, one doesn't really need to make reference to the standards of chemistry in affirming that water is a liquid, as this fact is quite visible. Secondly, someone's knowledge today that the sun is at the centre of the solar system likely makes no reference to (scientific) standards – it is simply a logical truth (a 'solar system' is a 'sun-centred system'). Moreover, Copernicus didn't believe so much that the sun is at the centre of the solar system, rather that it is the centre of the universe, and so he lacks knowledge not so much because of our higher standards today but simply because of his overall fundamental confusion about the structure of the universe.

Nevertheless, these critical points need not distract us from the basic insight, that in science – for many the best place to look for knowledge – one

[31] Peter Ludlow, "Contextualism and the New Linguistic Turn in Epistemology."

[32] Stanley, *Knowledge and Practical Interests*, 69.

usually finds references to epistemic standards in knowledge ascriptions. Following Ludlow, one might then anticipate that even "unembellished knowledge ascriptions, ones that do not contain explicit standards operators, nevertheless contain an unpronounced position for epistemic standards."[33] It is this suggestion that Stanley finds particularly objectionable. The basis for his concern is the observation (which he attributes to David Lewis) that 'standards talk' occurs in all sorts of discourses, not just in epistemic ones. Here Stanley cites the examples of:

1) By strict standards, France is not hexagonal.
2) By loose standards, this table is square.
3) By the standards of chemistry, what is in the Hudson River isn't water.[34]

He makes the point that, on the basis of the regularity with which ones finds these appended standards statements, one cannot "derive a conclusion about specifically epistemic context-sensitivity"[35] – and surely right about this, though this is not a derivation that Ludlow nor any standards contextualist needs to make. Rather, the relevant argument for the standards contextualist is normative: standards are regularly cited in scientific discourse, and since such discourse constitutes our best form of knowledge, standards should have a place in all forms of knowledge ascriptions. Be that as it may, one may follow Stanley's worry that, as motivated by the scientific model, "one would need standards positions in the syntax for virtually every predication,"[36] whether epistemic or not, a situation he describes as "deeply implausible." Moreover, he thinks the standards contextualist is committed to such an implausible conclusion, apparently because he sees the standards contextualist as arguing from the general ubiquity of standards discourse to its relevance to epistemic discourse – the argument we saw him also wrongly ascribing to standards contextualism above. But again this is not an argument a standards contextualist need subscribe to. There is no reason why standards contextualism should be committed to the claim that any sort of predication requires a standards position.

As Stanley describes the development of contextualism, starting with Fred Dretske's relevant alternatives theory, through to Gail Stine's contextualist improvement of Dretske's theory, and arriving at the versions of contextualism

[33] Stanley, *Knowledge and Practical Interests*, 69, italics removed.
[34] Stanley, *Knowledge and Practical Interests*, 70.
[35] Stanley, *Knowledge and Practical Interests*, 70.
[36] Stanley, *Knowledge and Practical Interests*, 71, italics removed.

formulated by Stewart Cohen and Keith DeRose,[37] one of the main motivations for the theory was to explain how one can know that one has hands, but not know that one is not a brain in a vat – that is, to explain the failure of deductive closure for knowledge. With standards contextualism, the explanation is essentially that the premise ("I know that I have hands") is understood to involve low standards for knowledge, whereas the conclusion ("I don't know that I am not a brain in a vat") involves high standards. Stanley submits that this contextualist interpretation of the case leaves its "oddity … unexplained."[38] But that is true only if one is committed to the failure of deductive closure in this case, and there is no reason why the contextualist should be so committed. There are in fact epistemologies that do require the failure of closure here, such as epistemologies that are committed to a sensitivity condition (i.e., one knows that p only if, were p not true, one would not believe p). But there is no necessity that contextualism be one of these epistemologies. Rather, a contextualist might suggest that the same low standards be used with both the premise and the conclusion, with the result that closure is preserved. So a contextualist can do better than explain the oddity – she can remove it.

This completes our responses on behalf of standards contextualism to the critiques offered in Stanley's *Knowledge and Practical Interests*. In their *Knowledge in an Uncertain World*, the practicalists Fantl and McGrath also raise objections to contextualism. We examine their arguments next.

3. Fantl and McGrath versus Contextualism

For Fantl and McGrath, it is fallibilism about knowledge – the thesis that one can know that p even though (one is aware that) there is a chance that p is false – that orients their discussion of the comparative merits of contextualism and practicalism. Once one gets over the puzzling nature of fallibilism, which Fantl and McGrath describe as the 'madness' of fallibilism,[39] one is left with the more practical question of determining how likely it must be that p is false for one's claim to knowledge to be withdrawn. In this regard they quote Laurence Bonjour, who comments that "it is … unclear what sort of basis or rationale there might be for fixing [this likelihood] in a non-arbitrary way."[40] Their answer to the problem

[37] Stanley, *Knowledge and Practical Interests*, 17-22.

[38] Stanley, *Knowledge and Practical Interests*, 72.

[39] Fantl and McGrath, *Knowledge in an Uncertain World*, 15.

[40] Laurence Bonjour, *Epistemology: Classic Problems and Contemporary Responses* (Lanham, MD: Rowman & Littlefield, 2002), 43, quoted in Fantl and McGrath, *Knowledge in an Uncertain World*, 25.

of what could constitute such a 'basis or rationale' is to point to what is practically at stake: p is probable enough to be known if it is "probable enough to be properly put to work as a basis for belief and action."[41] Fantl and McGrath don't say a great deal about what it means for a proposition to be 'properly put to work as a basis for belief and action,' but here is a familiar case that perhaps captures what they have in mind. The case, drawn from Rudner's 1953 paper "The Scientist *qua* Scientist Makes Value Judgments," concerns a scientist who, given a set of evidence, is considering the safety of a drug, and Rudner's claim is that the degree of confirmation of the hypothesis, 'this drug is safe,' is a function both of the (conditional) probability of this hypothesis given the evidence and "the importance, in the typically ethical sense, of making a mistake in accepting or rejecting the hypothesis."[42] As such, 'this drug is safe' is probable enough to be accepted for Rudner if the potential for harm from using the drug is low enough to be ethically acceptable, and this is arguably what Fantl and McGrath mean when they say that a proposition is "probable enough to be properly put to work as a basis for belief and action."

But there is a different way to answer Bonjour's challenge and find an alternative, non-arbitrary criterion that tells us when the probability of a proposition is high enough for this proposition to be the subject of knowledge. Fantl and McGrath assert that there is, at least, a probability that is low enough to categorically rule out a claim to knowledge: "presumably," they say, "there can't be knowledge with probability 0, probability 1/2, or even probability 2/3."[43] And to be sure, a probability of 1 would be high enough to ground a claim to knowledge. Given that we are stuck with fallibilism, is there then a probability less than 1 that can assuredly meet the challenge of knowledge? My suspicion is that reference to probabilities in an assessment of the evidential support for a hypothesis is unnecessary. Consider, for example, a piece of mundane perceptual knowledge, such as when someone knows that she sees a hand. In such a case, does a knower need to be aware of the probability of the truth of her belief given the available evidence? Typically, when someone knows that she has a hand, she will have no precise idea at all of how probable the truth of her belief is, short of its being 'high enough.' Rather, she will ground her knowledge on a variety of factors, such as the coherence of her belief with other beliefs she has, the evidence that her perceptual system is functioning properly (e.g., that she is not subject to

[41] Fantl and McGrath, *Knowledge in an Uncertain World*, 25.

[42] Richard Rudner, "The Scientist *qua* Scientist Makes Value Judgments", *Philosophy of Science* 20 (1953): 2.

[43] Fantl and McGrath, *Knowledge in an Uncertain World*, 25.

hallucinations, that her eyes are not diseased, and so on), and other matters whose impact on the justification of her belief is substantive, but cannot be usefully quantified. Once those factors have been settled upon and it is determined that they justify a belief, a knower may suggest that her belief has a strong likelihood of being true, without giving this likelihood a precise quantitative value. That is, probability assignments used in the justification of a claim are really just afterthoughts: they are ways of summarizing the quality of non-quantitative evidence. I think this is what we find in scientific contexts as well, with the exception being cases where statistical analysis is the core methodology.

What this means, then, is that Bonjour's challenge misconstrues the process of justification. Justifying a claim is not a matter of continually attempting to bump up the probability one attaches to this claim until a 'threshold' is reached. Such quantitative precision is usually not meaningfully attainable. And so the process of ascribing to oneself knowledge should not be viewed as a matter of deciding that one's belief is, in the first instance, 'probable enough' and on that basis justified. Typically the process is reversed: one ascribes to oneself knowledge and then asserts that one's belief has a high probability of being true (e.g., "I know that p, so it's probably true that p"). It follows that Fantl and McGrath themselves misconstrue the challenge facing the intellectualist, or as they call her, the 'purist.' They comment:

> the fallibilist who recoils at the thought of denying purism or allowing pragmatic encroachment should bear in mind her tasks: to explain away the apparent madness of fallibilism and to give us some idea of what it takes for a probability to be 'knowledge-level.' To retain purism (and deny pragmatic encroachment) she must perform these tasks without appealing to a conception of significant chances of error that allows stakes to play a role – that allows significance to vary without corresponding variance in your strength of epistemic position with respect to p.[44]

Fantl and McGrath are here contending that the purist, to cope with fallibilism, will need to find a probability level at which she can be said to know a proposition, and explain why this probability level has such an effect without making reference to practical matters. Moreover, since the strength of epistemic position is assumed to be fixed, she will not be able to explain the ability of this probability level to generate knowledge by pointing to a greater preponderance of evidence. This is a challenge that Fantl and McGrath do not think the purist can meet.

[44] Fantl and McGrath, *Knowledge in an Uncertain World*, 29.

But there is a conception of 'significant chances of error' that allows this significance to vary, without a corresponding variance in how much evidence the agent possesses. This is by means of a change in the epistemic standards that govern knowledge ascriptions. We should point out initially that by an 'epistemic standard' we do not mean that, in order to know a proposition, the probability that this proposition is true must reach a certain level. Again, epistemic standards primarily involve other matters than the probability that a proposition is true, since an agent typically lacks a well-grounded idea of the probability of a proposition (a key exception, again, are those sciences that make essential use of statistical methodologies). Rather, standards involve matters such as those mentioned above – e.g., the overall coherence of one's belief system and the well-functioning of one's perceptual apparatus – and a variety of other considerations that in many cases are unique to the subject matter. For instance, in the bank cases where the proposition of concern is "the bank will be open tomorrow (Saturday)," a fairly modest standard on the basis of which the agent could generate convincing evidence that the bank would be open tomorrow is whether the agent noticed that the bank was open the previous Saturday. By comparison, a more rigorous standard would require something like access to official bank policy detailing precise opening hours. Generally speaking, the modest standard works fine for normal day-to-day contexts, both where the practical stakes are low, but also where the practical stakes are not at all relevant. That is, where someone has no stake in the matter whether the bank is open (not just a low stake) the lower standard is perfectly acceptable in grounding a claim to knowledge. On the other hand, if the agent is a bank employee and she were asked whether the bank is open tomorrow, she would then be subject to a higher standard by virtue of her role. As such, it would be entirely unacceptable for her to make reference simply to the fact that the bank was open the previous Saturday in justifying the claim that the bank will be open tomorrow. Rather, she would need to make explicit reference to bank policy. Note that this need for a higher standard is not a result of the agent having higher stakes. There may in fact be nothing practically at stake for her in not using the higher standard, such as the threat of job loss. It's just that, as a bank employee, she should know better.

Ludlow's introduction of the notion of standards into a contextualist epistemology was motivated by scientific examples, and it is these examples that work best in illustrating how standards contextualism succeeds at answering the problem Fantl and McGrath pose for purism. Suppose we are asked, "Do we know that this table is brown?" Using the lower standards usual for quotidian life, one would look at the table under acceptable lighting, making sure one was not

examining only a covering for the table, and upon seeing a brown colour confidently ascribe to oneself knowledge that the table is brown. But of course fallibilism is true and one may be mistaken. For example, one may be subject to perceptual illusions where it can appear that a table is brown, though in fact it is not. The question Fantl and McGrath raise is, "When is the chance of error significant enough for one to retract one's (self-)ascription of knowledge?", to which they answer, "When the practical downside of being mistaken is high enough." The problem they pose for purism, then, is to provide an account of a 'significant chance of error' that does not refer to practical matters (while keeping the strength of one's epistemic position fixed), and here our answer is to suggest that in scientific contexts the standard for a table being brown is more rigorous than what we find in day-to-day contexts. In fact, strictly speaking, appearing to be brown is not good enough for a scientist to conclude that a table is brown since the appearance of brown is a product of one's psychology and not necessarily, truly representative of the colours of physical objects. Indeed, for a scientist, tables may not be brown nor any other colour for that matter (*qua* conglomerates of colourless atoms and molecules). So with a scientific standard, the claim to know that a table is brown is attended with a significant chance of error, and this conclusion is arrived at independently of any reference to practical stakes, and without having varied the strength of one's epistemic position (i.e., one perceptual evidence for the brownness of the table remains the same).

The situation in the bank cases is one where the agent knows that the bank will be open on Saturday when the stakes are low, but she lacks that knowledge when the stakes are high. It is assumed by most, contextualists and practicalists alike, that this is intuitively the correct interpretation of the situation. However, there is a sense in which this situation is somewhat problematic epistemically speaking, given that the strength of the agent's evidential position does not change in moving between the high and low stakes cases, given that she still believes the claim either way, and given that the truth value of the claim has not changed either. The general contextualist strategy is to suggest that the meaning of 'knows' varies in the two situations, and that it is this change of meaning that accounts for the differing epistemic assessments in the two cases. Yet as Stanley suggests,[45] and as echoed by Fantl and McGrath,[46] such a semantic strategy seems a bit too easy. Apart from the fact that changing the meaning of 'knows' can account for our intuitions in the bank cases, Stanley argues that "there is no further evidence that knowledge ascriptions are context-sensitive in a distinctively epistemological

[45] Stanley, *Knowledge and Practical Interests*, 155.

[46] Fantl and McGrath, *Knowledge in an Uncertain World*, 30.

way."[47] Similarly, Fantl and McGrath assert that in order for contextualism "to be plausible we need some independent evidence that the content of knowledge attributions can vary with speech context in ordinary non-philosophical contexts."[48] For them, as for Stanley, the capacity of contextualism simply to offer an interpretation of the bank cases is not enough to render contextualism convincing.

It is here that the reference to epistemic standards can make an important difference. It is sometimes claimed that scientific knowledge is the best form of knowledge one can obtain. The reason for this acclaim is the rigour with which scientists test their hypotheses. When a scientist tests a hypothesis using empirical evidence, strict standards are in place to ensure the accuracy of the empirical evidence as well as the cogency of the inductive step one takes from the evidence to the hypothesis under test. Thus, when a standards contextualist makes the point that knowledge ascriptions are context-sensitive in that they are relative to the standards in place in a particular context, there is in fact evidence that such ascriptions are context-sensitive in this distinctively epistemological way, evidence drawn from an observation of how knowledge claims are made in the sciences. Moreover, for the same reason, we thereby have independent evidence that the content of knowledge attributions do vary with speech context in ordinary non-philosophical contexts, specifically, when one moves from an everyday speech context to a scientific speech context where, as a norm of scientific discourse, one makes reference to higher epistemic standards in defending knowledge claims.

Having then responded to both Stanley's and Fantl and McGrath's objections to contextualism, let us conclude by examining two further, potential problems for, specifically, standards contextualism.

4. Two Problems for Standard Contextualism

To begin with, although the standards themselves make no reference to practical issues, one might argue that one's choice of standards is guided by practical matters, so there is no evading practicalism, after all. And in fact there is no doubt that practical issues can impel one to raise epistemic standards, such as in the Rudner case. However, one's motivation to choose an epistemic standard need not make any reference to practical issues but only to the requirement that a knowledge claim have an increased chance of being true. For example, this is the motivation of a scientist who adopts a rigorous standard leading to the conclusion

[47] Stanley, *Knowledge and Practical Interests*, 33.

[48] Fantl and McGrath, *Knowledge in an Uncertain World*, 30-31.

that a table isn't brown (or any colour for that matter). It's simply that, if we take into consideration atomic physics and the psychology of perception, it becomes apparent that the colours of objects are more a product or our minds than of what properties an object really has. The question of whether that fact has some sort of practical benefit is simply irrelevant. The same can be said for a bank employee who is asked whether her bank is open on Saturday. Here she uses the higher standard of referencing company policy, not because there is some practical benefit in doing, but only because this is the better, official way of finding out the truth as regards a bank's opening hours. Of course, the bank employee and the scientist could adopt a practical motivation, if they wish, for the conclusions they derive. They simply don't need to, and moreover it wouldn't sound right for either of them to cite practical benefits in justifying the standards they adopt. The scientist who says that she refers to atomic theory and the psychology of perception in answering questions about the colour of physical objects because she can make a profit by doing so should probably be distrusted on scientific issues. One should also be skeptical about the bank employee who cites company policy because it gives her a feeling of power. What would she say if that feeling of power led her in a different direction?

The second concern with standards contextualism is how one should justify the use of a lower epistemic standard, if a higher standard is available. Shouldn't one always defer to a higher epistemic standard in assessing knowledge ascriptions, if one's objective is unremittingly epistemic? It is important that the gravity of this problem for standards contextualism not be underestimated, for one aspect of contextualism is an element of equality in the various meanings that can be attached to the word 'know.' This equality stems from the fact that, from a contextualist perspective, many epistemic issues boil down to semantics. For example, in the philosophy classroom, one often uses the word 'know' in a strict way that leads to skepticism, on the assumption that there could be such things as brains in vats or Evil Demons. On the other hand, outside the classroom, the meaning of 'know' is more liberal and one need not guard against such extreme skeptical possibilities. It follows that, if the choice of the meaning of 'know' is guided primarily by epistemic concerns, one should likely use the higher (skeptical) standard found in philosophy classrooms. This problem has an analog when we turn to the epistemic standards used in everyday life as compared to those adopted in scientific practice. If a rigorous scientific perspective informs us that physical objects are not really coloured, and that colours are simply psychological constructs, then epistemically speaking everyone should believe that objects aren't coloured since that's the result of adopting the highest epistemic

standard. Of course, this is an unintuitive result since we do believe, and think we know, that objects are coloured (that, for example, a table is brown) in the normal course of daily affairs, despite the fact that this knowledge is grounded in the use of a lower epistemic standard.

In this regard, it's worthwhile pointing out that this problem does not arise if we are practicalists. Where there is not much is at stake in a knowledge ascription, a practicalist will condone the use of a lower epistemic standard. For example, whether the brownness of a table really inheres in a table or is simply a psychic construct, or whether I am a brain in a vat or a real human being, makes no practical difference if one is using the table for day-to-day uses. Thus, the practicalist has no trouble claiming to know that "this table is (really) brown" and "I am not a brain in a vat." Here we have an illustration of why Fantl and McGrath think practicalism to be so "attractive and easy," to be so "extremely plausible."[49] For them, the plausibility of practicalism "doesn't depend on the particular epistemic standards in force in the speech context" but is instead a result of a principle they call '*Action*,' that "if you know something which is relevant to your choice situation then you are proper to act on it."[50] Clearly, in normal contexts, it is proper for one to act on "this table is (really) brown" and "I am not a brain in a vat."

So, with standards contextualism, how do we justify the use of a lower epistemic standard, if a higher standard is available? The answer to this difficulty is to look more closely at how one goes about testing a hypothesis and, on the basis of these tests, judging that one is in a position to know the hypothesis. These tests will inevitably involve evidential reports regarding mundane observable objects whose existence and properties are largely taken for granted. That is, in assessing the reality of these mundane objects, a lower epistemic standard is taken and must be taken if the process of evaluating a hypothesis on the basis of a higher standard is to ever get started. One finds, paradoxically, such a preference occurring in the philosophy classroom, where philosophers debate the reality of mundane objects while feeling no awkwardness about taking the chairs they are sitting on and the pens they are using for granted. We also find this preference for lower standards in the psychology lab where the colours of objects are considered figmentary psychic constructs, but where researchers still calmly point out to each other apparently real colours of tables and chairs. In general terms, for an epistemic inquiry to proceed, one needs to work from a base of accepted claims, such as background assumptions and observational claims, whose evidential

[49] Fantl and McGrath, *Knowledge in an Uncertain World*, 51.

[50] Fantl and McGrath, *Knowledge in an Uncertain World*, 51.

support is held to a lower standard. This lower standard is needed so that the inquiry can gain deliberative traction: the philosophical discussion of skepticism needs to make commonplace assumptions about the world in order for the discussants to even communicate, and in the psychology lab if the researchers can't assume the intersubjective reality of the coloured, observed world, there would be no way for them to even discuss the psychic construction of the world. Thus even if we assume with contextualism that there is a preference for the use of higher standards regarding the use of the word 'know,' the practice of epistemic evaluation requires, nevertheless, a lower standard for the framing of evidential claims.

But one needs knowledge claims meeting lower standards not just for evidential assessments. Most knowledge claims are based on a background of other knowledge claims which must be accepted before the initial knowledge claim can itself be accepted. As an illustration, consider again the High Stakes bank case where there is much at risk for the agent if she fails to deposit her cheque on time. It is suggested that the agent doesn't know that the bank will be open on Saturday given the seriousness of missing a bill payment. Alternatively, the recommendation is that the agent should wait in the Friday line-up to deposit her cheque. But that recommendation holds only if there are other claims that we assume the agent knows in this case, claims that are relevant to her projected courses of action. For example, she may be mistaken about whether she's at the right bank, and so waiting in the Friday line-up could have disastrous results for her. Thus, the agent's reasoning, that she doesn't know that the bank will be open on Saturday and that she must go to the bank on Friday afternoon, relies on a previous knowledge claim that she knows this is the right bank. Now, the problem for the practicalist is this: given a high stakes, practical situation, it may be unreasonable to ascribe to the agent knowledge that she is at the right bank. In fact, given a high stakes, practical situation, it may be that the agent hardly knows anything at all relevant to the situation. So when arriving at the bank on Friday and seeing the long line-ups, what should the agent do? The analysis has now become highly complicated for the practicalist. With high enough stakes, the practicalist is soon driven to a practical and deliberative paralysis.

It is at this stage that standards contextualism shows its worth. As we have suggested, even where one has set high epistemic standards, the application of these standards requires the use of lower standards for knowledge claims that underpin the use of the higher standards. So lower standards will need to apply if we are to even get to use higher standards. But none of this makes sense for the practicalist: if practical circumstances compel one to retract one's knowledge

claims, even where the evidential support appears strong, then they will compel one to retract one's knowledge claims wherever they impact one's decision making. They do this because they force the agent to raise her epistemic standards both for knowledge claims under contention and for any relevant background knowledge claims: again, this is because the failure of any of these knowledge claims could have serious practical consequences. On the other hand, the procedure for the standards contextualist is to look to the standards themselves as a guide to whether one has knowledge. For any number of background claims and for a variety of evidential claims, one will adopt reduced standards simply as a way to move an investigation along. Now, even when a serious practical situation arises, these reduced standards will still apply, though the standard for the claim under contention may be heightened. This is because such claims will have wide application in a variety of areas of inquiry and decision making, and suspending them may be broadly disruptive. Nevertheless, there could be practical pressure to raise the standards for these claims and retract the presumption that one knows them to be true. At this stage, potential knowers will become conflicted: one knows these claims on a lower standard but doesn't know them on a higher standard. One finds this situation occurring in the philosophy classroom where skepticism is the topic of discussion and where, if we lack knowledge on anything at all, we lack knowledge that we're sitting in a philosophy classroom discussing skepticism. This conflict is resolved, as we all know, by abandoning the higher standard and allowing a lower standard. We do this to avert a practical and deliberative paralysis.

5. Conclusion

In this paper my goal has been to defend standards contextualism from criticisms posed against it by Jason Stanley, Jeremy Fantl and Matthew McGrath. Let me emphasize that in discussing epistemic standards, not just any standards will do. There are strict rules over what count as good epistemic standards, rules drawn from the quotidian norms on knowing and the heightened rules one finds in academia, science, professional fields and the like. I have not said much at all about what these particular standards are and how they are legitimated, only that they are directly relevant to the ascription of knowledge claims, and that practical matters have no bearing whatsoever on their legitimization. It has also been my contention that standards are not specifically designed to demand an increased probability for a claim to be warranted. Their use is usually much more qualitative than this. As such, the desire expressed by Bonjour and others to find a level of probability sufficient for knowledge is misconceived and even distracting. Finally,

we have with standards contextualism a reasonable and efficient analysis of the nature and legitimacy of skepticism: basically, skeptics are in the business of raising the standards on knowledge, an honorable and often worthwhile activity. Our response to skepticism is that, if we are not careful, standards will be raised to a point where epistemic investigation becomes impossible, and that situation is surely of not much practical benefit. Moreover, though it is good thing to know claims, one can often get along just fine (practically speaking) having only justified, true beliefs, or just true beliefs, or even just beliefs. For example, the agent in High Stakes may decide to wait in the Friday line-up after considering the risks of being wrong and wishing to be sure about her knowledge. But it may turn out that in doing this she is acting in an extremely impractical manner compared to a less cautious person who, despite her lack of knowledge, puts off the visit till the next day and finds the bank open with no line-ups.

INTERVENTIONISM DEFENDED

Kevin McCAIN

ABSTRACT: James Woodward's *Making Things Happen* presents the most fully developed version of a manipulability theory of causation. Although the 'interventionist' account of causation that Woodward defends in *Making Things Happen* has many admirable qualities, Michael Strevens argues that it has a fatal flaw. Strevens maintains that Woodward's interventionist account of causation renders facts about causation relative to an individual's perspective. In response to this charge, Woodward claims that although on his account X might be a relativized cause of Y relative to some perspective, this does not lead to the problematic relativity that Strevens claims. Roughly, Woodward argues this is so because if X is a relativized cause of Y with respect to some perspective, then X is a cause of Y simpliciter. So, the truth of whether X is a cause of Y is not relative to one's perspective. Strevens counters by arguing that Woodward's response fails because relativized causation is not monotonic. In this paper I argue that Strevens' argument that relativized causation is not monotonic is unsound.

KEYWORDS: causation, intervention, manipulation, James Woodward, Michael Strevens

A commonsense way of thinking of causal relationships is that they are relationships that can allow one to bring about changes through various manipulations. In other words, if X causes Y, then manipulating X in appropriate ways should lead to changes in Y. This commonsense intuition lies at the heart of manipulability theories of causation.[1] In order for a manipulability theory to have any hope of being an acceptable theory of causation it must provide an account what counts as the appropriate ways of manipulating X. A promising account of appropriate ways of manipulating X is James Woodward's account, which defines the appropriate ways of manipulating as 'interventions.' According to Woodward, "an intervention I on X with respect to Y will be such that I causes a change in X, I does not cause a change in Y via some route that does not go through X, and I is exogenous in the sense of not itself having a cause that affects Y via a route that does not go through X."[2] Woodward's interventionist account of causation

[1] See, for example, R.G. Collingwood, *An Essay on Metaphysics* (Oxford: Clarendon Press, 1940), Georg von Wright, *Explanation and Understanding* (Ithaca, NY: Cornell University Press, 1971), Peter Menzies and Huw Price, "Causation as a Secondary Quality," *British Journal for the Philosophy of Science* 44 (1993): 187-203, and James Woodward, *Making Things Happen* (Oxford: Oxford University Press, 2003).

[2] James Woodward, "Cause and Explanation in Psychiatry: An Interventionist Perspective," in *Philosophical Issues in Psychiatry: Explanation, Phenomenology and Nosology*, eds. K. Kendler

(hereafter referred to simply as 'interventionism') offers a powerful tool for understanding the nature of various causal relations that fits with the commonsense view of the connection between causation and manipulation.

Although Woodward's case in favor of interventionism is persuasive, interventionism is not without its detractors.[3] Recently, Michael Strevens has attacked interventionism on the grounds that it introduces a problematic relativity to facts about causation.[4] More precisely, Strevens argues that interventionism is committed to the claim that the facts concerning whether X is a contributing cause of Y are dependent upon one's perspective because contributing causation is defined with respect to a variable set. In response to this charge, Woodward attempts to provide a de-relativized notion of contributing causation.[5] Woodward's response involves distinguishing between being represented as a contributing cause (following Strevens this will be referred to as 'relativized causation') and being a contributing cause simpliciter.[6] Woodward claims that although X is a relativized cause of Y relative to a particular variable set, it does not lead to the relativity that Strevens claims. Roughly, Woodward argues this is so because if X is a relativized cause of Y with respect to some variable set $\mathbf{V}$, then X is a contributing cause of Y simpliciter. Woodward's argument rests on the assumption that relativized causation is monotonic in the sense that adding variables to $\mathbf{V}$ will not lead to X's no longer being a relativized cause of Y.[7] Strevens concedes that if successful, Woodward's response would provide a de-relativized notion of contributing causation.[8] However, Strevens argues that Woodward's response fails because relativized causation is not monotonic in this

---

and J. Parnas (Baltimore, MD: Johns Hopkins University Press, 2008), 139. See Woodward, *Making Things Happen*, 98 for a more precise formal definition of 'intervention.'

[3] Since my purpose here is to defend interventionism from a particular objection rather than provide a full-scale argument for its acceptance, the interested reader is encouraged to consult Woodward, *Making Things Happen*, for a thorough presentation and defense of interventionism.

[4] Michael Strevens, "Essay Review of Woodward, *Making Things Happen,*" *Philosophy and Phenomenological Research* 74 (2007): 233-49.

[5] James Woodward, "Response to Strevens," *Philosophy and Phenomenological Research* 78 (2008): 193-212.

[6] Michael Strevens, "Comments on Woodward, *Making Things Happen*," *Philosophy and Phenomenological Research* 78 (2008): 171-92.

[7] When Woodward and Strevens speak of adding variables to a variable set, they do not mean that the phenomenon 'in the world' that is being represented is changed in any way. Adding variables to a variable set amounts to giving a more detailed account of the phenomenon in question.

[8] Strevens, "Comments on Woodward."

way. Given the assumed failure of Woodward's response, Strevens concludes that interventionism has the unacceptable consequence of entailing that causation is relative to one's perspective.

My goal in this paper is to defend interventionism by demonstrating that Strevens' argument is unsound. In section one, I lay out the relevant aspects of interventionism as defined by Woodward in *Making Things Happen (MTH)*. Additionally, I describe how Woodward modifies the notion of contributing cause in response to Strevens' initial objection. Also, I explain how this modification to the notion of contributing causation is supposed to meet Strevens' challenge of providing a de-relativized notion of contributing causation. In section two, I explicate Strevens' argument against the monotonicity of relativized causation. I also describe both the example that Strevens uses to support a key premise in his argument and the example that he uses to motivate his overall argument. In the third and final section, I argue for the falsity of a key premise in Strevens' argument. In addition, I argue that the example that Strevens presents to support this premise is problematic. What is more, in this section I also explicate why the example that Strevens uses to motivate his overall argument is flawed.

1. Interventionism and De-Relativized Contributing Causation

In order to appreciate Woodward's attempt to provide a de-relativized notion of contributing causation as well as the moves made in the dialectical exchange between Woodward and Strevens, it is necessary to be clear about the precise definitions of the following notions: *intervention*, *direct cause*, and *contributing cause*. To begin, the general idea of an intervention is fairly straightforward. Interventions are manipulations upon one or more variables in a system under idealized experimental conditions. An intervention on a variable X should be understood in terms of experimental manipulations of X that are well designed for determining if X causes Y in an idealized experimental setting (an experimental setting that excludes confounding influences). More precisely, "Is assuming some value $I=z_i$, is an intervention on X with respect to Y if and only if I is an intervention variable for X with respect to Y and $I= z_i$ is an actual cause of the value taken by X."[9] Woodward explains that I is an intervention variable on X with respect to Y just in case:

(**IV**) I1. I causes X

I2. I acts as a switch for all the other variables that cause X. That is, certain values of I are such that when I attains those values, X ceases to

[9] Woodward, *Making Things Happen*, 98.

> depend on the values of other variables that cause *X* and instead depends only on the value taken by *I*.
>
> I3. Any directed path from *I* to *Y* goes through *X*. That is, *I* does not directly cause *Y* and is not the cause of any causes of *Y* that are distinct from *X* except, of course, for those causes of *Y*, if any, that are built into the *I*-*X*-*Y* connection itself: that is, except for (a) any causes of *Y* that are effects of *X* (i.e., variables that are causally between *X* and *Y*) and (b) any causes of *Y* that are between *I* and *X* and have no effect on *Y* independently of *X*.
>
> I4. *I* is (statistically) independent of any variable *Z* that causes *Y* and that is on a directed path that does not go through *X*.[10]

Simply put, the idea is that an intervention on *X* is some sort of change that an experimenter in an ideal setting can bring about in *X* that is such that the method of bringing about that change will directly affect only *X* and the method of changing *X* will exclusively set the value of *X*.

The next notion that needs to be defined is *direct cause*. According to Woodward, *X* is a direct cause of *Y* with respect to variable set **V** if and only if there is "a possible intervention on *X* that will change *Y* (or the probability distribution of *Y*) when all other variables in **V** besides *X* and *Y* are held fixed at some value by interventions."[11]

Now the notion that lies at the heart of the debate between Woodward and Strevens, *contributing cause*, needs to be examined. Woodward defines a contributing cause in the following manner:

> A necessary and sufficient condition for *X* to be a (type-level) contributing cause of *Y* with respect to variable set **V** is that (i) there be a directed path from *X* to *Y* such that each link in this path is a direct causal relationship; that is, a set of variables $Z_1 \ldots Z_n$ such that *X* is a direct cause of Z_1, which is in turn a direct cause of Z_2, which is a direct cause of $\ldots Z_n$, which is a direct cause of *Y*, and that (ii) there be some intervention on *X* that will change *Y* when all other variables in **V** that are not on this path are fixed at some value. If there is only one path *P* from *X* to *Y* or if the only alternative path from *X* to *Y* besides *P* contains no intermediate variables (i.e., is direct), then *X* is a contributing cause of *Y* as long as there is some intervention on *X* that will change the value of *Y*, for some values of the other variables in **V**.[12]

---

[10] Woodward, *Making Things Happen*, 98. Woodward notes that any unqualified instance of 'cause' in this definition should be understood to mean contributing cause.

[11] Woodward, *Making Things Happen*, 55.

[12] Woodward, *Making Things Happen*, 59.

Strevens argues that interventionism makes causation relative to one's perspective because according to Woodward's definition, whether or not X is a contributing cause of Y is a relative matter.[13] He supports this claim by pointing out that contributing cause is defined with respect to a variable set. So, Strevens challenges Woodward to explain how interventionism can provide an account of contributing causation that is not dependent upon one's perspective. That is, Strevens challenges Woodward to de-relativize the notion of contributing cause.

In light of this criticism and the challenge put forward by Strevens, Woodward makes the following modifications. First, he claims that the above definition of contributing cause from *MTH* would have been better put as necessary and sufficient conditions for "X to be correctly represented as a contributing cause of Y with respect to $\mathbf{V}$."[14] Second, he claims that "One can then go on to say that X is a contributing cause of Y *simpliciter* (in a sense that isn't relativized to any particular variable set $\mathbf{V}$) as long as it is true that there exists a variable set $\mathbf{V}$ such that X is correctly represented as a contributing cause of Y with respect to $\mathbf{V}$."[15]

Given Woodward's modifications it is easy to see how he responds to Strevens' challenge. First, the original notion of contributing cause is not to be understood as a definition of contributing causation, but instead as a definition of representation as a contributing cause (relativized causation). Second, Woodward offers a de-relativization of contributing causation by claiming that X is a contributing cause of Y (in a non-relativized sense) if and only if X is a relativized cause of Y relative to some variable set. As Strevens notes, Woodward's de-relativization of contributing causation is a success only if relativized causation is monotonic.[16] That is, Woodward has defended interventionism from the charge of relativity only if it is true that "if X is a relative cause of Y with respect to a variable set V, then it is also a relative cause of Y with respect to any superset of V."[17]

Now that the relevant items of interventionism have been defined and Woodward's response to Strevens' challenge explained, I will explicate Strevens' argument against the monotonicity of relativized causation.

---

[13] Strevens, "Essay Review of Woodward."
[14] Woodward, "Response to Strevens," 209.
[15] Woodward, "Response to Strevens," 209.
[16] Strevens, "Comments on Woodward."
[17] Strevens, "Comments on Woodward," 175.

2. Strevens' Attack on Monotonicity

Strevens argues that Woodward's attempt to provide an unrelativized notion of contributing causation fails because relativized causation is not monotonic. More specifically, Strevens argues that X may be a relativized cause of Y relative to variable set $\mathbf{V}$, but not a relativized cause of Y relative to variable set $\mathbf{V^*}$ (a variable set constructed by adding more variables to $\mathbf{V}$). Strevens offers the following formulation of his argument:

> 1. Adding variables to a variable set can sometimes make relativized causal relations appear (as monotonicity allows).
>
> 2. A variable's counting as an intervener depends on the *non-existence* of certain relations of relativized causation.
>
> 3. Thus (from (1) and (2)), variables may lose their status as interveners as other variables are added to the variable set.
>
> 4. A variable's status as a relativized cause requires the existence of an intervener with respect to which a certain further condition is satisfied. If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes.
>
> 5. Thus (from (3) and (4)), variables may lose their status as relativized causes as other variables are added to the variable set.[18]

Concerning line one, Strevens correctly notes that it is consistent with monotonicity. One can easily see how this premise is true given interventionism. To illustrate this, consider a variable set, $\mathbf{V}$, which includes only X and Y where X is a relative cause of Y. Suppose further that as a matter of fact there are intermediate causal links between X and Y (this is an unproblematic supposition because it is plausible that there are intermediate links between any causally related variables – excepting, perhaps, those that are representative of features of fundamental physical reality). If the variable set $\mathbf{V}$ were supplemented with one of these intermediate causal links, Z, relativized causal relations will appear in the augmented variable set $\mathbf{V^*}$. Specifically, relativized causal relations between Z and both X and Y that did not hold relative to $\mathbf{V}$ will hold relative to $\mathbf{V^*}$.[19, 20]

---

[18] Strevens, "Comments on Woodward," 175-76.

[19] Strevens also appeals to an example to help support this premise. However, it is not necessary to describe this example in detail because the example is problematic and, further, it is not needed to support this premise.

[20] Michael Strevens has mentioned (personal correspondence) that this is not quite how he intended to support this premise. Strevens thinks that this premise is true because he thinks that adding variables to a variable set, $\mathbf{V}$, can lead to the appearance of relativized causal relations

Line two of this argument is true largely by definition. According to Woodward, *I*'s being an intervention on *X* with respect to *Y* requires that *I* not be a contributing cause of *Y* via some causal path that does not go through *X*. Strevens points out that the notion of an intervention does not explicitly appeal to relativization to a variable set, but it does appeal to the notion of contributing causation. He correctly notes that the notion of contributing cause that is appealed to in the definition of an intervention is either relativized or not. If it is relativized, then this premise is true. If it is not relativized, then it not clear that Woodward's attempt to provide an unrelativized notion of contributing causation is successful. As Strevens says "it is far from clear that an account of unrelativized causation that takes the form of a definition invoking unrelativized causal facts constitutes a genuine derelativization."[21] While it is not certain that an unrelativized notion of causation cannot be crafted in this way, Strevens is correct in claiming that it is not clear that it can. So, the truth of this premise should be granted.

Strevens supports his premise in line four by way of an example. In Strevens' example an experimenter is interested in determining whether bottled water consumption is a cause of heart disease. According to Strevens, a bungling experimenter may manipulate subjects' bottled water consumption by increasing their intake of salty foods. Since eating salty foods increases one's chance of heart disease, there will be a correlation between increased bottled water consumption and heart disease in this case. Strevens claims that relative to the variable set that only includes bottled water and heart disease; salty food intake will count as an intervention on bottled water consumption with respect to heart disease. So, relative to this variable set bottled water consumption will count as a cause of heart disease. However, Strevens maintains that if we augment this variable set with other variables, such as artery hardening, a relativized causal relation between salty food intake and heart disease will appear; thus, revoking the status of salty food intake as an intervention on bottled water consumption. As a result of the failure of salty food intake to count as an intervention relative to this variable set, bottled water consumption will lose its status as a relativized cause of heart disease.

---

between the members of the original set **V**. For example, Strevens thinks that even though *X* and *Y* are causally unrelated relative to **V**, adding variables to **V** can make a relativized causal relation between *X* and *Y* appear. I think that Strevens is mistaken on this point, however, since Strevens and I both agree that this premise is true, the issue of what makes this premise is true can be set aside for the moment.

[21] Strevens, "Comments on Woodward," 181.

Although Strevens uses similar examples to support the premises in lines one and four of his argument, he notes that these two examples each involve a different set of assumptions. So, in an effort to bolster his overall argument Strevens presents a third example, which purports to describe a situation in which both premise one and premise four are true. Here is Strevens' third example. Again he is considering the relationship between drinking bottled water (*B*) and heart disease (*H*). In this example Strevens stipulates that eating salty foods (*S*) sometimes leads one to drink bottled water and sometimes leads one to drink red wine (*W*). He assumes that if eating salty foods leads one in a particular instance to drink bottled water, then it does not lead to her drinking red wine; and vice versa. He also stipulates that this sporadic consumption of red wine is enough to off-set the effect that eating salty foods has on heart disease by hardening one's arteries (*A*). Strevens points out that eating salty foods may be used to manipulate bottled water consumption, but it will only be successful some of the time because eating salty foods sometimes leads to red wine consumption. However, Strevens claims that in the subset of cases where eating salty foods successfully manipulates bottled water consumption there will be a correlation between bottled water consumption and heart disease because there will be no red wine consumption to off-set the effect eating salty foods has on heart disease. He goes on to say that the fact that *S* is not a legitimate intervention on *B* with respect to *H* will only show up when the variable set under consideration includes either *W* or *A*. So, Strevens maintains that *B* will be a relativized cause of *H* relative to the variable set {*B, H, S*}. However, since S will not be an intervention on *B* with respect to *H* relative to variable sets {*B, H, S, W*}, {*B, H, S, A*}, or {*B, H, S, W, A*}, *B* will not be a relativized cause of *H* relative to any of these variable sets. Strevens argues that his example illustrates a situation where *B* is a relativized cause of *H* relative to a particular variable set, but *B* fails to be a relativized cause of *H* relative to variable sets constructed by adding further variables to the original variable set. Thus, Strevens concludes that this example reveals that relativized causation is not monotonic.

3. In Defense of Monotonicity

My strategy for defending Woodward's claim that relativized causation is monotonic from Strevens' attack is straightforward. I argue that the premise in line four of Strevens' argument, "If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes," is false. First, I argue that the example that Strevens appeals to in defending this premise fails to provide evidence for its truth. Second, I explicate why Strevens' third example, which is

designed to motivate his overall argument, is problematic. Third, I explain why in general arguments of the kind Strevens offers fail to establish that relativized causation is not monotonic.

In order to demonstrate that the premise "If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes" is false, it is important to first spend a little time re-examining the notion of an intervention. Specifically, it is necessary to examine the conditions under which a variable *I* counts as an intervention on *X* with respect to *Y*. For the sake of simplicity, in the course of evaluating these conditions I will consider a very sparse variable set, **V**, which includes only *I*, *X*, and *Y*. According to Woodward's definition *I* is an intervention on *X* with respect to *Y* just in case *I* is an actual cause of the value taken by *X* and *I* is an intervention variable for *X*. In order to be an intervention variable for *X*, *I* must satisfy four further conditions of **IV**: 1) *I* has to cause *X*, 2) *I* has to act as a switch for all other variables that cause *X*, 3) *I* must not be a direct cause of *Y* nor a cause of any causes of *Y* that are distinct from the causal connection *I*-*X*-*Y*, and 4) *I* must be statistically independent of any variables that cause *Y* without causing *X*. Utilizing the sparse variable set mentioned above {*I, X, Y*}, let us assume that *I* is an intervention on *X* with respect to *Y*. So, *I*'s having the value that it does causes *X* to have its actual value and *I* is solely responsible for *X* having that value (this follows from the first condition of an intervention and the first two conditions of **IV**). Additionally, there is no possible intervention on *I* that will change the value of *Y* when *X* is held fixed at a certain value via other interventions (this follows from condition three of **IV** and the definition of direct cause). Finally, there are no causes of *Y* that are also causes of *I* (this follows from condition four of **IV**).

At this point I will turn to a critical assessment of the example Strevens uses to support the premise "If a variable loses its status as an intervener, then, other variables may lose their status as relativized causes." Recall that Strevens' example involves the incompetent experimenter who uses salty food intake (*S*) to manipulate bottled water consumption (*B*) in order to determine if bottle water consumption (*B*) causes heart disease (*H*). Strevens maintains that if we simply consider the variable set {*S, B, H*}, *S* will be an intervention on *B* with respect to *H*. So, *B* will be a relativized cause of *H* relative to this variable set because using *S* to intervene on *B* will lead to a change in *H*. However, he claims that if further variables such as artery hardening (*A*) were added to this variable set, *S* would lose its status as an intervention. Thus, *B* would not be a relativized cause of *H* relative to the augmented set {*S, B, H, A*}.

This example is ineffectual because, contrary to what Strevens claims, *S* is not an intervention on *B* with respect to *H* even in the impoverished three variable set. The reason that this is the case is obvious. *S* fails to meet the third condition of **IV**. That is, relative to the variable set containing {*S, B, H*} *S* will be a direct cause of *H*. *S* is a direct cause of *H* relative to this variable set because there are possible interventions on *S* that lead to changes in *H* while *B* is held fixed. Keeping with Strevens' example, one such intervention would be for the experimenter to lock the subjects in a room with only salty food to eat while providing the subjects a fixed quantity of bottled water to drink and nothing else. If the experimenter intervenes on salty food intake while holding bottled water consumption fixed in this manner, there will be a correlation between salty food and heart disease.

One might worry that this way of responding to Strevens' example sneaks in an unrelativized notion of causation into the notion of an intervention. The concern here is that whether *S* is a direct cause of *H* relative to this variable set depends on relations between *S*, *B*, *H*, and variables outside of the variable set under consideration. One might think that what is really occurring here is an illegitimate appeal to facts about all of the variables there are for determining what counts as an intervention. So, the worry is that appealing to these facts about variables outside of the variable set under consideration is utilizing unrelativized causation.[22]

There are two ways of responding to this worry. The first way to respond is to point out that there is nothing in this response to Strevens' example that appeals to an unrelativized notion of causation. Direct causation is a relativized notion of causation. Further, appealing to the existence of variables that can possibly be added to a variable set is not invoking unrelativized causation. The second way to respond is to draw attention to the fact that the satisfaction of the third condition of **IV** only requires that it not be possible to manipulate *S* in a way that will affect *H* while *B* is held fixed. Again nothing about this suggests an illicit appeal to unrelativized notions of causation.

[22] This objection bears some similarity to an unrelated objection that Strevens, "Essay Review of Woodward," raises for Woodward's account of causation. Strevens argues that in order to determine whether *I* is an intervention we need to know about the causal relations that *I* bears to other variables. In order to determine these relations we need to intervene on *I*, so we need another intervention variable *I** that is an intervention on *I*. However, in order to determine whether *I** is an intervention we need to know about the causal relations that *I** bears to other variables, and so on. Thus, there seems to be a problem with ever determining whether *I* is an intervention.

Given my treatment of the example Strevens uses to support this premise, the problem with the example he uses to motivate his overall argument is probably apparent. However, at the risk of a bit redundancy I will briefly explain what is wrong with this example as well. Remember that in this example Strevens stipulates that eating salty foods (*S*) sometimes leads one to drink bottled water (*B*) and sometimes leads one to drink red wine (*W*). He assumes that if eating salty foods leads one in a particular instance to drink bottled water, then it does not lead to her drinking red wine; and vice versa. Further this occasional consumption of red wine is enough to off-set the effect that eating salty foods has on heart disease (*H*) by hardening one's arteries (*A*). Strevens claims that in the subset of cases where eating salty foods successfully manipulates bottled water consumption there will be a correlation between bottled water consumption and heart disease because there will be no red wine consumption to off-set the effect eating salty foods has on heart disease. He believes that the fact that *S* is not a legitimate intervention on *B* with respect to *H* will only show up when the variable set under consideration includes either *W* or *A*. So, Strevens maintains that *B* will be a relativized cause of *H* relative to the variable set {*B, H, S*}. However, since S will not be an intervention on *B* with respect to *H* relative to variable sets {*B, H, S, W*}, {*B, H, S, A*}, or {*B, H, S, W, A*}, *B* will not be a relativized cause of *H* relative to any of these variable sets.

This example shares the same problem as the previous example. Namely, *S* does not meet the conditions for being an intervention on *B* with respect to *H*. Again in this example *S* is a direct cause of *H* relative to this variable set because there are possible interventions on *S* that lead to changes in *H* while *B* is held fixed. One such intervention would be the one I described above, giving someone only salty food to eat and a limited supply of bottled water to drink and nothing else. In this situation changes in *S* would be correlated with changes in *H* while *B* is held fixed. Since *S* is not an intervention on *B* with respect to *H*, *B* will not be a relativized cause of *H* relative to the variable set {*S, B, H*}. So, Strevens' example fails to illustrate a situation where a variable that is a relativized cause relative to a variable set ceases to be so when more variables are added to the variable set.

One might try to defend Strevens' example from my objection by having *S* represent not only the eating of salty food, but instead the eating of salty food in an environment where bottled water and red wine are both freely available. If *S* represents the eating of salty foods in an environment where bottled water and red wine are both freely available, one may think that the experiment that I describe above will not count as an intervention on *S*. The idea is that given what *S* represents in this case, an experiment that only gives someone salty food to eat

and bottled water to drink will not be intervening on *S* because the experiment requires an environment that differs from the one specified by *S*. So, the experiment that I describe would fail to be an intervention on *S* with respect to *H*. Since the experiment does not intervene on *S*, it cannot show that *S* is a direct cause of *H*. Thus, the experiment described above cannot demonstrate that *S* is not an intervention on *B* with respect to *H*, and so, it fails to pose a problem for Strevens' example.[23]

Although it is true that understanding *S* as representing eating salty foods in the specified environment will make it the case that the experiment I describe is not an intervention on *S* with respect to *H*, this move will not save Strevens' example. The problem for Strevens' example does not go away by having *S* specify a particular environment. Even if *S* represents eating salty foods in an environment where bottled water and red wine are both freely available, there are possible interventions on *S* that lead to changes in *H* while *B* is held fixed. One such intervention would be to perform an experiment where subjects are given only salty food to eat and they are forced to drink so much bottled water that although there is red wine available, they will not drink any. This experiment will count as an intervention on *S* and it has the same result as the original experiment that I described, namely, it results in subjects who eat salty foods and only drink bottled water. So, this experiment will expose the causal connection between *S* and *H* because changes in the amount of salty food eaten in this circumstance will lead to changes in heart disease, and hence, this experiment demonstrates that *S* is not an intervention on *B* with respect to *H*. Thus, even if *S* represents eating salty foods in a specified environment, Strevens' example is still problematic.

Now that I have shown why Strevens' argument is unsuccessful I will explain why other arguments of this kind will also fail to demonstrate that relativized causation is not monotonic. As noted above, there are very specific conditions under which *I* may be properly said to be an intervention on *X* with respect to *Y*. Given an understanding of these conditions, it is possible to see how adding variables to variable set **V**, which includes *{I, X, Y}*, might result in *I*'s no longer meeting those conditions. Since *I* is the sole cause of *X* having the value that it does and *I, X, Y* are the only variables in the set, *I* will be a direct cause of *X* relative to this set. Adding variables to **V** can make it the case that *I* is no longer a direct cause of *X*. For instance, there may be a variable *Z* that is a causal link between *I* and *X*. However, this fact will not mean that *I* is no longer an intervention because adding variables will not make it the case that *I* fails to be a cause of *X*. So, adding variables will not lead to *I*'s failing to satisfy the first

[23] Thanks to Michael Strevens for drawing my attention to this objection.

condition for being an intervention nor the first condition for being an intervention variable. The second condition of **IV** requires I to essentially sever the causal links between X and all other variables besides Y. It seems possible that if enough variables are added to **V**, there will be some for which I fails to function as a switch. So, it seems possible that adding variables to **V** may lead to I failing to meet the second condition of intervention variables. Since I is not a direct cause of Y relative to **V**, then adding variables to the set will not lead to I becoming one. Likewise, adding variables to the set will not lead to I becoming a cause of Y along another path than the I-X-Y path. So, adding variables will not lead to I's failing to satisfy condition three of **IV**. Finally, if I is statistically independent of any cause of Y (as condition four requires), then adding further variables to **V** will not lead to I and Y having a common cause. So, it seems that the only way that adding variables to a variable set can remove I's status as an intervention is by introducing variables that are causally related to X that are such that I cannot block their causal impact on X.

In order for arguments like Strevens' to be successful it has to be the case that there are ways that a variable can lose its status as an intervener which result in other variables losing their status as relative causes. However, the only way that I can lose its status as an intervener by adding variables to **V** is for it to fail to disrupt the causal connections between some of these new variables and X. This situation will not result in variables losing their status as relativized causes in general and it will not result in X losing its status as a relativized cause of Y in particular. So, an evaluation of the conditions for I being an intervener shows that Strevens' argument and other arguments of the same kind cannot demonstrate that relativized causation is not monotonic.

Assuming that what I have said here is correct, Strevens fails to show that relativized causation is not monotonic. More generally, arguments of the kind that Strevens presents cannot establish that relativized causation is not monotonic. Thus, it reasonable to conclude that Woodward's attempt to provide an unrelativized notion of contributing causation is successful. Which in turn means that it is reasonable to think that interventionism does not render causation relative in the problematic way that Strevens suggests.[24]

[24] Thanks to Michael Strevens and Jim Woodward for helpful discussion of this topic. A very special thanks to Brad Weslake for comments on numerous earlier drafts and for many fruitful discussions of these issues.

ON THE EPISTEMOLOGY OF MODAL RATIONALISM: THE MAIN PROBLEMS AND THEIR SIGNIFICANCE[1]

Mihai RUSU

ABSTRACT: In this paper, I discuss the main characteristics of the epistemology of modal rationalism by proceeding from the critical investigation of Peacocke's theory of modality. I build on arguments by Crispin Wright and Sonia Roca-Royes, which are generalised and supplemented by further analysis, in order to show that principle-based accounts have little prospects of succeeding in their task of providing an integrated account of the metaphysics and the epistemology of modality. I argue that it is unlikely that we will able to develop an exhaustive and accurate principle-based account that discriminates objectively between correct and deviant modal knowledge. Even if such an account can be formulated, a non-circular way of justifying its necessity also seems to be out of our reach.

KEYWORDS: metaphysical necessity, modal epistemology, modal rationalism, principle-based account, Christopher Peacocke

1. Introduction

The modal scepticism of early analytic philosophers is largely explainable by the traditional commitment of analytic philosophy to empiricist epistemology. According to the classical empiricist view about modality,[2] there are two types of true propositions: those concerning knowledge of the 'external world' (what Hume calls 'matters of fact') and a priori truths (logical, mathematical, and semantic knowledge, what Hume would call 'relations of ideas').[3] Necessity is the exclusive attribute of the latter type, precisely due to lack of factual content. The reliability of factual knowledge is ultimately grounded in a causal relation between our senses and the objects of experience, but any proposition that is causally grounded cannot be necessary, as no matter how many particular empirical inputs confirm it, there is no way one could ward off theoretically any

[1] This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/159/1.5/S/133675.

[2] See David Hume, *An Enquiry Concerning Human Understanding* (Oxford: Oxford University Press, 2007), 18.

[3] Hume discusses only mathematical truths as cases of knowledge of 'relations of ideas.' The (Kantian) extension of the a priori to analytic and logical truths is, however, uncontroversial.

exception (as we do in mathematical proofs, for instance). The acknowledgement of necessary factual truth is therefore precluded by a traditional empiricist commitment.

The famous case set forth by Kripke turned the tides in the analytic tradition in favour of the recognition of the existence of necessary a posteriori truths.[4] What needs to be stressed here is that this reversal is more profound than a simple reconsideration of modal notions: one cannot accept that there are necessary factual truths and remain committed to classical empiricism, on pain of incoherence. One of the available alternatives is to endorse a rationalist viewpoint. Kripke and like-minded philosophers[5] are *rationalists* concerning modal knowledge, that is, they hold that some modal knowledge is a priori and, moreover, that *a posteriori modal knowledge is dependent on a priori modal knowledge*. Also, rationalists hold that modal truth is mind-independent – more precisely, they uphold *a substantive account of modal truth* and maintain that *modal knowledge latches onto mind-independent content*. In fact, the central problem of modal epistemology – how do we have modal knowledge? – is truly meaningful (and pressing) in accounts that maintain that there is a distinct type of mind-independent modal knowledge; and rationalist theories are typical examples of this sort of account.[6] There is no problem of modal knowledge for empiricist / naturalist sceptics concerning modality: knowledge of modal truths is typically reduced to knowledge of a priori truths. If the a priori is regarded as problematic, then so is the modal, but there is no special difficulty regarding modal knowledge beyond whatever problem the a priori leads to. Also, at the other end of the spectrum, there is no epistemological problem if modal knowledge is seen as no different from ordinary empirical knowledge (or if ordinary empirical knowledge is deemed to have a built-in modal component).[7] Even if a slight difference of character between ordinary empirical knowledge and modal knowledge is

[4] Saul A. Kripke, *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1980).

[5] See Stephen K. McLeod, "Rationalism and Modal Knowledge," *Crítica* 41 (2009): 32 for a comprehensive list of doctrines regarding modality, including names of prominent modal rationalists.

[6] The need for an epistemological account also appears for those who hold that modal truth is literally reducible to truth in possible worlds, as in the case of Lewis' theory of possible worlds in David K. Lewis, *On the Plurality of Worlds* (Oxford: Blackwell, 1986). The lack of a causal relation with other worlds, which the possible-world theorist posits as *truly* existing, generates an analogue of Benacerraf's dilemma for the case of modal realism.

[7] See Crawford Elder, "An Epistemological Defence of Realism About Necessity," *Philosophical Quarterly* 42 (1992): 317-336; and Nenad Miščević, "Explanining Modal Intuition," *Acta Analytica* 18 (2003): 5-41.

recognised, the troublesome character of modality can be explained away if our modality-grasping ability is considered an extension of a natural ability of human beings, namely, our ability to evaluate counterfactuals.[8]

The central problem of modal epistemology arises accordingly only when the theorist regards ordinary truth and knowledge, on the one side, and modal truth and knowledge, on the other, as substantially different. The epistemological programme of modal rationalism is then correctly characterized as the attempt to bridge the gap between ordinary knowledge and modal knowledge by proceeding from the content of grounding a priori modal knowledge. In this paper, I attempt to formulate some sceptical concerns regarding rationalist attempts of transcending the properly acknowledged distance between ordinary truth and modal truth. I will not consider the details of all modal rationalist theories, but instead I will focus on one typical development of this doctrine, to wit, Peacocke's *moderate rationalism*. I build on criticism of Peacocke's theory by Wright and Roca-Royes and show that some of the critical insights of these authors, supplemented by other related concerns, can be extended to all forms of modal rationalism. But before exploring the significance of the mentioned criticisms, a brief exposition of Peacocke's account of modality and its underlying purpose is required.[9]

2. Peacocke's Account of Metaphysical Modality

Peacocke understands quite clearly the acute need of harmonising the metaphysics and the epistemology of modality and considers it a focal point of his endeavour. He calls the task of reconciling the account of the content of our true statements (metaphysics) with an account of how we are able to know that content (epistemology) the Integration Challenge.[10] His *Being Known* is dedicated to developing accounts that meet 'the Integration Challenge' not only for necessity, but also for other important philosophical notions, such as freedom, self-knowledge and intentional content, and the past. Peacocke maintains that the case of metaphysical necessity cannot be elucidated by means of a causal epistemology,

[8] See Timothy Williamson, *The Philosophy of Philosophy* (Oxford: Blackwell, 2007); Boris Kment, "Counterfactuals and Explanation," *Mind* 115 (2006): 261-310; and Boris Kment, *Modality and Explanatory Reasoning* (Oxford: Oxford University Press, 2014) for theories based on this claim.

[9] The summarisation of Peacocke's conception of modality follows Peacocke's own brief account of his ideas in Christopher Peacocke, "Summary," *Philosophical Books* 42 (2001): 81-83, and Roca-Royes' exposition in Sonia Roca-Royes, "Modal Epistemology, Modal Concepts and the Integration Challenge," *Dialectica* 64 (2010): 335-361.

[10] Christopher Peacocke, *Being Known* (Oxford: Oxford University Press, 1999), 1.

but rather by using a principle-based account, as modal truth is "fundamentally an a priori matter."[11] This is the basic content of Peacocke's rationalist position concerning metaphysical modality: there are implicitly known a priori principles that determine the truth-value of modal statements. The truth of a priori statements is derived from the understanding conditions (and ensuing determination theories) of their constituent expressions.[12] Peacocke's declared goal is to provide a plausible intermediary way between Lewis' modal realism and mind-dependent accounts.[13] In this regard, he holds that there is a mind-independent component to modal truth, but also that the metaphysical investigation of the modal domain should be akin to Strawson's descriptive metaphysics, that is, it should proceed from the structure of our thought, in this case, of modal thought.[14]

An assignment s has, in Peacocke's account, the form of a quadruple <D, val, propval, ext>, composed of: the domain of objects D; the function val – the semantic value of concept C – which assigns an extension (object, truth-value, etc.) to C; the function propval – the property value of C – which assigns a property or a relation to (atomic concept) C; and ext – the extension – which assigns extensions to properties and relations. Every assignment has a corresponding specification (s-specification), which is the set of thoughts that the assignment counts as true. Then,

> A specification is a genuine possibility iff there is some admissible assignment which counts all its members as true.[15]

That is, if a set of thoughts is to count as representing a real or genuine possibility, it must be true in an admissible assignment. Admissibility is defined at the level of assignments as compliance with all the Principles of Possibility – the implicitly known a priori principles that bear upon our modal knowledge. The Principles of Possibility are divided by Peacocke into three main categories.

The first category includes only one principle, called the Modal Extension Principle (MEP), which can be stated in the following form:

> **MEP**: An assignment s is admissible only if: for any concept C, the semantic value of C according to s is the result of applying the same rule as is applied in the determination of the actual semantic value of C.[16]

[11] Peacocke, "Summary," 82.
[12] Peacocke, *Being Known*, 143.
[13] Cf. Peacocke, "Summary," 82.
[14] Peacocke, *Being Known*, 2.
[15] Peacocke, *Being Known*, 126.

To take one of Peacocke's examples, for the concept <bachelor>, an assignment will be admissible only if the semantic value that it assigns to this concept is obtained as the intersection of the extension of <man> according to s and the extension of <unmarried> according to s. While the extension of a concept may vary in different possible worlds (i.e. possible specifications, admissible assignments), the limits of its variation are fixed by the rules that determine its semantic value in the actual world.

The Modal Extension Principle is complemented by a class of Constitutive Principles which function at the level of objects, properties and relations, whereas MEP works at the level of concepts and sense.[17] The overall significance of Constitutive Principles is the following: an assignment counts as admissible (thereby determining a possible specification) "only if it respects what is constitutive of the objects, properties and relations it mentions."[18] Forthwith, Peacocke propounds some candidates that could plausibly fit the role of Constitutive Principles. One of his examples concerns the fundamental kind of an object:

> **Fundamental Kind:** If P is a property which is an object x's fundamental kind, then an assignment is inadmissible if it counts the proposition x is P as false.[19]

Necessity of origin is also discussed as a plausible constitutive principle. But what is important to note here is that Peacocke doesn't attempt to provide an exhaustive list of constitutive principles (not even at a highly general level or for some significant types of objects, e.g. a list of what should count as constitutive principles for all objects, living beings, artefacts, etc.). Nor does Peacocke argue that the few principles he does consider are something more than plausible variants for Constitutive Principles, to wit, they must be actual implicitly known a priori principles shared by an overwhelming majority of the community of language users. In fact, he explicitly states that establishing which principles are true is not his main concern, but rather emphasising the role of the Principles of Possibility in our understanding of modality and developing a general framework for further investigation of the domain.[20] This important point will be developed further on in the paper.

---

[16] Peacocke, *Being Known*, 136.
[17] Peacocke, *Being Known*, 144.
[18] Peacocke, *Being Known*, 144.
[19] Peacocke, *Being Known*, 145.
[20] Cf. Peacocke, *Being Known*, 191.

Finally, there is also a second-order[21] plenitude principle, called the Principle of Constrained Recombination (PCR):

> **PCR:** An assignment is admissible if it respects all the previous principles (MEP and the Constitutive Principles).[22]

PCR states that MEP and the Constitutive Principles are jointly sufficient for admissibility. As such, any statement included in a specification that is represented by an admissible assignment will be judged as possible.

The Principles of Possibility determine what counts as possible from a metaphysical point of view – whatever is true in an admissible assignment (an assignment that respects the principles) is deemed genuinely possible. But they also provide an account of our modal knowledge because they are taken to constitute the possession conditions for the concept <possible>: if one doesn't have implicit knowledge of the principles, then one doesn't possess the concept <possible>. But when one has the concept, one also has the means to determine if some specification is genuinely possible, to wit, one has epistemic access to metaphysical modality.[23] This guarantees, according to Peacocke, that there is no divorce between the metaphysical and the epistemic aspects of modality – they are by these means connected, and the Integration Challenge is met. It is important to note, in relation to this point, that MEP is recursive, as it applies to the concept <admissible> itself (and thereby to the concept <possible>). MEP and the other principles make up the rule for <admissible>, therefore determining its actual extension. But MEP can be applied to the concept <admissible> itself – if something is to count as admissible, it must respect the actual rules for <admissible>.[24] One important consequence of the self-applicability of MEP is that our characterization of necessity will be itself necessary. The gist of the argument is pretty straightforward.[25] Necessity is defined as truth in all the admissible assignments. But this rule of necessity can become an object for MEP – therefore, under every admissible assignment s, the semantic value of <necessary> will include all and only the thoughts which are true under every admissible assignment according to s. But then the characterisation of necessity will be true in every admissible assignment, i.e. necessary. This is as it should be if modal truth is taken to be mind-independent – if our conception of necessity were contingent, and thereby necessary truths were only contingently necessary, then there could

[21] A second-order principle is a principle that makes reference to other principles.

[22] Cf. Peacocke, *Being Known*, 149.

[23] Cf. Roca-Royes, "Modal Epistemology, Modal Concepts," 339.

[24] Cf. Peacocke, *Being Known*, 151.

[25] Cf. Peacocke, *Being Known*, 152.

be multiple equally entitled characterisations of necessity. This would lead to an admission of the possibility of incompatible necessary statements, and thereby to the conclusion that a strong notion of metaphysical necessity is untenable. Therefore, necessary truths must be necessarily so in any realist account, if that account is to succeed.

3. Wright's Criticism

In a paper about Peacocke's theory of necessity,[26] Crispin Wright highlights three structural limitations of Peacocke's principle-based account. I will only dwell upon the first structural problem singled out by Wright.[27]

Wright's argument starts from Peacocke's remark that MEP is recursive, which supposedly allows Peacocke to show that his characterisation of necessity is itself necessary. Wright points out, rightfully in my opinion, that there is a difference between the principles being true under every admissible assignment (according to the given account) and something being metaphysically necessary. For the account to work, one already has to know that Peacocke's characterisation of necessity is necessarily correct, more precisely, metaphysically necessarily correct. Peacocke's account has to characterise the right type of necessity (true metaphysical necessity) and not some other notion (Peacocke-necessity, as Wright puts it). For the characterisation of necessity to be necessary according to Peacocke, it does not have to be the right one; whatever the Constitutive Principles are taken to be, if MEP stays in place, the characterisation will come out necessary, i. e. true in all admissible assignments. So, it has to be determined independently if the characterisation of necessity is the right one, to wit, that the way we construe admissibility by means of Constitutive Principles is correct.

Wright maintains that this is a symptom of a structural challenge that lies at the heart of Peacocke's account, and is encapsulated in the admission that there is a distance between metaphysical necessity and our knowledge of it, viz., that there is an Integration Challenge concerning metaphysical modality. "Accepting that challenge is accepting that we need to integrate a satisfactory account of the constitution of necessity with a satisfactory account of its epistemology,"[28] Wright argues, but whatever that integrated account is, it will exhibit the same problem. According to Peacocke's account, what one can do is just find a purportedly constitutive property, and then attempt to give an account of how we are able to

[26] Crispin Wright, "On Knowing What is Necessary: Three Limitations of Peacocke's Account," *Philosophy and Phenomenological Research* 64 (2002): 655-662.
[27] Cf. Wright, "On Knowing What is Necessary," 656-659.
[28] Wright, "On Knowing What is Necessary," 658.

recognise that property. But if the constitutive part is not itself recognised as necessary, the account is compromised. Instead of explaining our knowledge of necessity, the epistemological part can only provide an account of our knowledge of constitutive properties. In light of these remarks, Wright concludes that "[t]he success of the account thus depends upon our recognition of a necessity which it cannot itself explain."[29]

If Wright's contentions are correct, the structural problem affects a whole class of accounts of modality, not only Peacocke's theory. Any reductive account of modality that proceeds by way of mind-independent constitutive properties encounters the problem of justifying its own necessity. In case our characterisation of necessity is contingent, we are strongly entitled to doubt whether what we recognise as constitutive/essential properties that generate necessary truths are indeed so, or this is rather a mind-dependent matter. And this brings us to the heart of the matter: the necessity of the account must already be acknowledged for it to proceed. But then knowledge of necessity is left unexplained. In Peacocke's case, the recursive character of MEP might be a sign of its running-on-any-fuel nature: whatever we 'pump' into it, the account still works. Even if this is not so, it still shows the inescapable requirement of explaining the impetus of the account. How do we recognise that our knowledge of the constitutive is the correct one? Peacocke's answer would be that this is an a priori matter – we have an implicit knowledge of constitutive principles. And this brings us to the re-statement of an important point: this structural doubt affects especially (perhaps exclusively) rationalist theories, that is, specific inquiries that proceed from necessary a priori principles of modal knowledge, but at the same time maintain that the modal domain is fundamentally mind-independent. The principles that ground modal knowledge must themselves be metaphysically necessary, but in virtue of what are they so? Even if the possession conditions for the concept <necessity> are correctly stated, they cannot provide the metaphysical explanation of the necessity of principles. If they could, there wouldn't be any Integration Challenge to consider. But then, if we uphold the necessity of the principles by way of constitutive facts, we need an independent integrated (metaphysical and epistemological) account of this grounding. What is it that makes the principles necessary and how do we know it? Whatever answer we concoct, this new characterisation must itself be necessary, so we seem to be left with a potentially infinite regress. Potentially infinite because the account can

[29] Wright, "On Knowing What is Necessary," 659.

stop at a certain point in its back and forth necessity-essentiality[30] movement and we may decide that necessity is the primitive notion. For instance, one may submit that the Principles of Possibility are just necessary, but we do not (cannot) have knowledge of the constitutive facts that ground them. This is not wrong, but it is a straightforward admission of defeat for the modal Integration Challenge.

Are modal rationalist accounts bound to fail because of this structural problem? Lowe endorses such an answer, but in relation to all accounts that aim to reach mind-independent truth by proceeding from the content of our concepts and words:

> [T]he fundamental mistake is to suppose […] that […] an 'advance' would have to proceed from a basis in our knowledge of our concepts and words – that is, from a knowledge of how we conceive of and describe the world – to a knowledge of that world 'as it is in itself,' independently of our conceptual schemes and languages. This 'inside-out' account of how knowledge of mind-independent reality is to be acquired already makes such knowledge impossible and must therefore be rejected as incoherent.[31]

But before I discuss the impasse of rationalist accounts of a purported mind-independent modal reality, I will consider a complex criticism of Peacocke's theory by Roca-Royes that I believe to be related to Wright's argument.

4. Roca-Royes on Peacocke's Principle-Based Account

I will summarise here Roca-Royes' criticism of Peacocke's account, insisting on what I regard as its most significant aspects.

Roca-Royes notes that the epistemological problem is not solved simply by providing a principle-based account of modality, but rather transferred from the modal domain (where it is explicit, according to Roca-Royes) to the constitutive domain (the essential properties that are encapsulated in the Constitutive Principles, which Peacocke holds that we know implicitly).[32] Consequently, the task of providing an epistemology of the constitutive is urgent for Peacocke. The role of an epistemology of the constitutive would then be to propound a procedure by means of which we attain explicit knowledge of (the correct) Constitutive

[30] In this context, I take 'essence' to be interchangeable with 'constitutive fact', *viz.*, whatever grounds modal truth without being itself modal. Peacocke's Constitutive Principles aim to reveal precisely such a grounding reality.

[31] E. Jonathan Lowe, "Essentialism, Metaphysical Realism, and the Errors of Conceptualism," *Philosophia Scientiae* 12 (2008): 28.

[32] Cf. Roca-Royes, "Modal Epistemology, Modal Concepts," 340-341.

Principles. Peacocke sketches a corresponding solution in a subsequent paper.[33] He argues that we reach explicit knowledge of the constitutive by using a two-step abductive process. At first, we identify some a priori known modal propositions (the proposition that if a living being originates in gametes a and b, then it necessarily originates in gametes a and b would be an appropriate example of such an a priori modal truth) that we use as the abductive base. Then, we search for the best explanation for the meaning of necessity that would be in accordance with the truth of the a priori modal propositions. The best explanation is, according to Peacocke, that necessity conforms to the Principles of Possibility (taking again the example given above, necessity would conform to this instance of Essentiality of Origin). Roca-Royes argues that the appeal to the a priori known modal propositions is ineffective, as all modal knowledge should depend upon the Principles of Possibility. When, as in the first step of Peacocke's abduction, the Principles haven't yet been established, there can be no claim to warranted modal knowledge. So, an appeal to independent (regarding the Principles of Possibility) route is required.[34]

To support the argument, Roca-Royes describes two cases where the abductive process would not yield the right kind of explicit constitutive knowledge: one where we have only implicit false beliefs about the constitutive realm, and the second where it is by mere epistemic luck that the concepts constituting our modal knowledge track mind-independent essential truth. In the first case, we would arrive at false explicit modal beliefs, and in the second at true modal beliefs, but that would not amount to knowledge. To conclude, the account is dependent on the correctness of our implicit beliefs about the constitutive realm, but even if they are correct, a full-fledged positive account of knowledge of the constitutive and how precisely it aligns with mind-independent facts is still required. Roca-Royes elaborates on her arguments by noting that we can develop accounts for concepts very similar to the purportedly correct one, and they all characterize some (potentially interesting) property (e.g., logical possibility, conceptual possibility, natural possibility, but also indefinitely many others). This raises the question of how we achieve the correct concept among so many (slightly or less slightly) deviant ones. The question in its turn emphasizes the fact that we need a full-fledged account of our knowledge of the constitutive.[35]

---

[33] Christopher Peacocke, "The Past, Necessity, Externalism and Entitlement," *Philosophical Books* 42 (2001): 106-117.

[34] Roca-Royes, "Modal Epistemology, Modal Concepts," 341-342.

[35] Cf. Roca-Royes, "Modal Epistemology, Modal Concepts," 342.

Roca-Royes contrasts Peacocke's set of rules for possibility (MEP + the Constitutive Principles) with another account that adds the following rule to MEP, instead of the Constitutive Principles:

> **(Const)** An assignment s is admissible only if, for any entities $e_1 \ldots e_n$ (objects, properties or relations) and for any n-ary relation R, $n \geq 1$, such that entities $e_1 \ldots e_n$ constitutively stand in the relation R, s does not count $Re_1 \ldots e_n$ as false.[36]

The alternative account cannot provide a non-conditional characterisation of possibility. Roca-Royes uses this difference to note that Peacocke conveniently builds just the right amount of content into the modal concepts so that his moderate rationalism works (the implicitly known Constitutive Principles cover all potential cases). In contrast to this, with the alternative definition one needs an independent knowledge of constitutive facts in order to ascertain the possibility of something. However, Peacocke's account is just as ineffective, because the Constitutive Principles are not argued for by using an independent characterisation.[37] Moreover, Roca-Royes suggests that the possession conditions for modal concepts are too demanding because they provide a full-fledged theory about the constitutive realm, which we supposedly possess implicitly. Modal disagreement is a further reason to doubt the appropriateness of Peacocke's account, as is Peacocke's allowing that there are principles that are unknowable explicitly.[38]

This concludes my summary of Roca-Royes' critical examination of Peacocke's moderate rationalism. All is now in place for a reflection on the significance of the criticisms explored above. In contrast to Roca-Royes and more in line with Wright's suggestions, I take these difficulties to be a symptom of a profound vulnerability of rationalist accounts of modal notions in general, and not only of Peacocke's account. In order to argue for this point, I will attempt to radicalize Wright's and Roca-Royes' arguments in order to extend them to a wider class of philosophical theories. Some supplementary arguments will be formulated. This will mark a clear departure from Roca-Royes' realism and her ultimate epistemological optimism concerning metaphysical modality.[39]

---

[36] Roca-Royes, "Modal Epistemology, Modal Concepts," 353.

[37] Cf. Roca-Royes, "Modal Epistemology, Modal Concepts," 354-355.

[38] Cf. Roca-Royes, "Modal Epistemology, Modal Concepts," 355-357.

[39] Stated at Roca-Royes, "Modal Epistemology, Modal Concepts," 336-337.

5. The Quandaries of Modal Rationalism

There are two levels of inquiry of modal rationalist theories that are both plagued by problems, as seen in the counterarguments examined above. I will now consider the two levels separately and show where Wright's and Roca-Royes' arguments can be supplemented or radicalized in order to affect a wider class of accounts.

a. The level of content

The level of content is constituted, naturally, by the specific metaphysical and epistemological accounts of real modality. First, we have the problem of:

a*. The need for a complete account

The need for a complete list of the a priori principles that determine our modal knowledge is inescapable. In the case of Peacocke's account, this requirement is pressing because the content of his Constitutive Principles encodes the possession conditions (hence the actual content) of modal concepts. In the absence of a complete statement of principles, both the characterisation of our concepts and the metaphysical description of mind-independent modal truth are ineffective (not only the theory has a spectral object, but it is itself spectral). Actually, this is a problem that affects all modal rationalist doctrines that are not fleshed out in an exhaustive account. To my knowledge, there is no attempt to provide such a complete account to date. Typically, modal rationalist accounts are only sketched, as in the cases of Kripke's and Peacocke's work. General guidelines are given, and suggestions for plausible a priori principles are adduced, but the needed statement of principles remains fragmentary and disparate. This way of handling the issue may prove a strategic advantage – if less a priori principles are endorsed, the chances that the account is disputed are smaller. Theoretically, the correct account of the constitutive and the characterisation of our modal knowledge may diverge. If, as in Peacocke's case, the metaphysical explanation of the constitutive realm is reflected exactly in our modal concepts, the strategic benefits of not attempting to provide the full account of modality are all the more obvious. The problem would be not only that the given account of the constitutive realm is wrong, but also that the conditions that are imposed on our notions and our knowledge are misapprehended.

These remarks on the need of a complete and accurate integrated account of metaphysical modality are not all there is to the story. A problem that is not so obvious has to do with the thesis that our (grounding) a priori modal knowledge is

implicit. Again, this is not an idea that is to be found exclusively in Peacocke's account. Modal rationalists would want to hold (albeit not always explicitly, one may quip) that a priori principles are implicitly known. Remember that in the introduction I elaborated on the reasons of developing a modal rationalist account. The Integration Challenge is central to rationalist thinking about metaphysical modality – rationalists agree that there is a gap between the metaphysical and the epistemological aspect of modality. The gap is explained, but also explained away by the claim that we have implicit knowledge of the Constitutive Principles – we seem to have trouble assembling the two perspectives, because our knowledge is not explicit; nevertheless, we can ultimately do it, because our implicit knowledge is still knowledge, and can be made explicit with some theoretical effort. But if not all our implicit knowledge is made explicit, we cannot pretend to know some fundamental facts about our conception of modality. For instance, we don't even know if our conception is consistent. We are entitled to believe that our modal knowledge (including our grounding modal knowledge) is vast, so that it can cover indefinitely many possible situations. If some of it remains hidden, then there is no way of knowing precisely that it doesn't contain incompatible principles. The fact that we have conflicting intuitions about modality in different situations makes this worry powerful enough.[40]

a**. The 'just the right amount' objection

Roca-Royes takes issue with Peacocke's building just the right amount of content into modal concepts. I hold that this is unavoidable if (any form of) modal rationalism is to play an adequate explanatory role. Now, a correct statement of the main stance of modal rationalism is that all a posteriori modal knowledge is dependent on some a priori modal knowledge. The modal force of a posteriori modal knowledge is transferred from the a priori principles. What one wants then is that these principles are effective, but also that there is no exception to the rule, viz., that there is no a posteriori modal knowledge that is not dependent on a priori knowledge. So, for this to happen, our a priori principles have to cover all cases of a posteriori modal knowledge. This puts modal rationalism in a very uncomfortable epistemological position. If modal truth is mind-independent, one may want to hold that there are possibilities that we don't know of (perhaps even cannot know of), for we don't have knowledge of all that is real and even in the cases of things we do know, we sometimes lack knowledge of all that is constitutive. The problem that our theory of the possible may be disproved by

[40] Theseus' Paradox would be an example of a case where intuitions rejecting the principle of Essentiality of Composition collide with some other intuitions that appear to support it.

further discoveries is meaningful, but hardly disconcerting. Its significance lies exactly in assisting our understanding of how much in our modal concepts is actually thinker-dependent (revising our view of possibility would always proceed at first by revising our non-modal concepts). But this is not what should bother us here.

The troublesome aspect of modal epistemology is revealed by accepting that the conditions on our possession of concepts are too demanding (as Roca-Royes seems to hold). But if these conditions are too demanding, then at least some Constitutive Principles don't do the work that is required of them. This means that there are some known modal propositions that are not obtained in the rationalist way (by being deduced from a priori principles plus empirical information regarding the possession of a certain property). But if there are pieces of modal knowledge that are not grounded by any specific a priori principle, then why should we think that we need a priori principles for modal knowledge at all? It is only natural to suppose that if there is a shorter route to modal knowledge, and, moreover, if some of our modal knowledge is not grounded by the a priori, then the shorter route is the right one.

To recap some earlier insights, the need for rationalism presents itself only when one acknowledges that there is, in Peacocke's terms, an Integration Challenge concerning metaphysical modality, to wit, we lack an epistemological account that explains adequately how we know modal facts. This gap is supposedly filled in by a priori principles, but these principles are required to be effective and all-encompassing. The gist of modal rationalism is that a priori content grounds and explains modal knowledge. But if some modal statements are not grounded in our concepts (contrary to what the modal rationalist holds), then the rationalist theory lacks actual explanatory power, and the Integration Challenge is not met. There still are modal truths that lack a corresponding epistemology. The question remains: how do we know these modal truths to be necessary/possible?

There is also a related point that we can make. It is clear that Roca-Royes' minimal principle (Const) is ineffective independently, as it is conditional upon previous knowledge or, alternatively, on a sceptical or agnostic stance regarding metaphysical modality (and serves all these accounts indiscriminately). But we should also notice that such a principle is indispensable in every type of reductive account of modality, so every account is, at least minimally, a principle-based account. If one believes that essence grounds metaphysical modality (and, correspondingly, knowledge of essence/knowledge of constitutive properties grounds modal knowledge), then one needs (Const) to link the essential and the

modal. However, even though it is plausibly a priori, (Const) is not a grounding principle of modality. A posteriori modal knowledge is not dependent on (Const), but rather on knowledge of constitutive properties and relations.

a***. Modal disagreement

The greatest challenge for modal rationalist accounts is, in my opinion, brought by the problem of modal disagreement. Roca-Royes takes modal disagreement to be a powerful argument against our modal concepts being as rich in content as Peacocke would want it. My point is related, but different. Suppose that two philosophers develop two different and incompatible integrated accounts of modal knowledge, but in accordance with the general character of rationalist theories (that is exhibited in Peacocke's theory) and without disagreeing on the actual empirical facts that are true. For instance, one holds that Essentiality of Origin is a universal Principle of Possibility, and one holds that it is not. Now, for Peacocke, one of them is making a mistake or is not in possession of the modal concepts. But this leads to a dilemma, as there is no principled way to decide who is right. Each one may adduce equally powerful independent reasons for her point of view. There may be no manifest incoherence in their doctrines. So what can we say about this situation? To my mind, the only way to decide between the two, all the while respecting the general principles of modal rationalism (mind-independent modal truth combined with a priori dependence of our modal knowledge), is to find an objective mind-independent criterion that would settle matters. One account should lead to correct modal knowledge, and one not. It is, however, very difficult to see what such a criterion might be. I cannot find any particular fact that would help us decide who is right and who is wrong. The reader should remember that the two philosophers may agree on all ordinary non-modal facts, but disagree regarding our modal statements, to wit, they disagree on the limits we impose on characterising other possible situations. These limits are not given by the empirical facts themselves, but by our stance on what counts as constitutive. The facts themselves may be mind-independent, but characterising them as essential seems to be thinker-dependent. Providing an epistemology of essence, in whatever guise, means most of all explaining how it is that some facts impose themselves on our knowledge as being essential, and others do not. It is my contention that rationalism cannot provide an adequate epistemology of the constitutive, and the reason for this has a lot to do with modal disagreement. All the empirical facts being acknowledged to be the same, two informed and penetrating thinkers may disagree, as they often do, about the principles of possibility. Whatever may settle the dispute (if this can happen) could only look

like a decision to (re)characterise things in a certain way. If there is no fact of the matter about which account is the correct one, then this shows that our take on what is to count as essential (or as a principle of possibility) is a mind-dependent affair.

It is very interesting that, from another perspective, modal disagreement casts serious doubt on the a priori character of the Principles of Possibility, if we take the a priori to be structural. It is implausible that the principles are actually being thought of in rationalist accounts as akin to Kripke's examples of contingent a priori truths.[41] Even if this were so, they would be deprived of the modal force that they are held to imbue modal a posteriori knowledge with. The principles also seem to have nothing of the highly abstract and sophisticated character of some mathematical notions that sometimes impends us from ascertaining mathematical truth. Yet, consensus seems hard to reach regarding the truth of modal principles. Could it be then that the principles are just expressions of decisions on what ordinary facts are to be held fixed when entertaining counterfactual hypotheses? Modal disagreement, supported all the more by the lack of any complete principle-based account of modality, points us toward an affirmative answer to this question.

A cumulative conclusion of my discussion so far is that Peacocke doesn't provide us with the much needed solution of discriminating between the correct and the deviant accounts of metaphysical modality. I argued that we have serious reasons to be pessimistic about the perspectives of formulating a satisfying principle-based account. This will become all the more obvious in the following discussion of the meta-level of principle-based rationalist accounts.

b. The meta-level

I have shown that some of the arguments that can be raised against Peacocke's theory (the most interesting ones, to be sure) are extendable to a whole class of accounts, namely, to modal rationalist accounts. If this is correct, then the fault must originate not in the misgivings of particular content, but rather in metatheoretical aspects that underpin Peacocke's arguments and theses. The explanation of the quandaries of moderate rationalism has therefore less to do with Peacocke's theorising, which is actually quite ingenious, and more to do with the epistemology of philosophy, to wit, with the general characteristics of the philosophical perspective that is assumed (modal rationalism, in this case). It is these structural metatheoretical aspects that lead to certain solutions being

[41] The idea of a contingent a priori truth is itself controversial.

formulated and employed.[42] So, it is not at all mistaken to suggest that what I have described in this section as being problematic in relation to the content of theories actually has an important metatheoretical component. More precisely, the difficulties that are related to the elaboration and development of theoretical content are plausibly caused by certain higher-order characteristics.

The reason that I am discussing the following difficulty under the 'meta-level' heading (when in fact all problems I treated in this chapter are in a significant way metatheoretical) is that this one is metatheoretical par excellence. The problem arises for an account (e. g., Peacocke's) in its entirety (it doesn't regard just one specific part of it, as the other problems do) and, if I am right, it is a general problem that affects every type of principle-based account (and thereby all forms of modal rationalism).

This second-order difficulty has already been described in an informal manner at the end of the section that was focused on Crispin Wright's criticism. I will now restate it in a more argumentative form. But first I will recap how the problem is made explicit in Wright's and Roca-Royes' criticisms.

In my opinion, Wright and Roca-Royes both detect the same problem regarding Peacocke's delineation of necessity and describe it in similar terms. Wright argues that the recursive character of MEP is useless for grounding the necessity of Peacocke's own characterization of necessity if we don't already know that the account singles out the right kind of modal notion (metaphysical necessity) and not some other similar concept. But this is exactly what is in need of justification in Peacocke's account, so no prior grasp of necessity can be invoked. Due to its self-applicable nature, Peacocke's definition of necessity works with every noetic fuel – necessary truth will turn out true in every possible specification, but this cannot guarantee by itself that we are employing the correct notion of necessity. As we have seen, Roca-Royes uses a very similar strategy when she disputes Peacocke's claim that knowledge of the principles is arrived at by means of an abduction that proceeds from some modal propositions that are known a priori. If modal status is grounded in and inherited from the principles, there can be no warranted claim to modal knowledge without the principles being already established. Wright also criticizes the overall strategy of explaining the modal by means of the constitutive. He claims that this type of account manages at best to give an epistemological explanation of the reductive notion, but leaves knowledge of necessity unexplained (although the account requires recognition of

[42] See Gary Gutting, *What Philosophers Know: Case Studies in Recent Analytic Philosophy* (Cambridge: Cambridge University Press, 2009) for more on the epistemology of philosophy and the two levels of theory building.

necessity to put the theoretical machinery in motion). Now, this last worry can be addressed in a fairly satisfying way by using something like Roca-Royes' (Const) principle, viz., something that links (knowledge of) modality with (knowledge of) essence in an appropriate way.[43] The epistemology of necessity is thereby reduced to the epistemology of essence. But all this talk of the need for prior knowledge of necessity or about the absence of an adequate modal epistemology may obliterate the real character of the predicament of principle-based accounts and their deep structural vulnerability. So, let me explain this vulnerability by describing the content of the metatheoretical problem in a question-and-answer form.

The fundamental question is the following:

Q1: Is our[44] characterisation of necessity contingent or necessary?

A1: The characterisation of necessity is contingent.

A1 is unacceptable for someone who (like Peacocke and all modal rationalists) holds that modal truth is mind-independent. If we allow the characterisation of necessity to be contingent, then this can very plausibly be linked to the fact that modal truth is fundamentally mind-dependent. The reader must note that this is not a claim that the fact that someone arrives at (the correct) characterisation of necessity is contingent, in the same way that the fact that we developed mathematics is contingent, but mathematical truths are nevertheless necessary. A1 encapsulates the thesis that the very definition of necessity (and the corresponding principles that it comprises) are contingent, so there may be different equally entitled characterisations of purported metaphysical necessity.

So, the modal rationalist can only endorse:

A2: The characterisation of necessity is necessary.

But in order to uphold A2, the modal rationalist cannot appeal to the necessity of first-order constitutive principles, as the principles are part of the characterisation that needs justification. A recursive principle like MEP will not do, as long as the correct real necessity of the account is not established. Therefore, the modal rationalist needs a corresponding second-order principle. So, now we have:

Q2: In virtue of what is the characterisation of necessity necessary?

[43] It is open for discussion if (Const) manages to do that, but this is not my concern here. I only cite (Const) as an example of how an explanatory link between the constitutive and the modal should be and can be established.

[44] Where 'our' refers to any principle-based characterisation.

The proponent of a principle-based account can only give an answer along the lines of:

> A3: The characterisation of necessity is necessary in virtue of some constitutive fact(s) that is/are encoded in one/several Constitutive Principle(s).

Then, the problem reappears. The new question is:

> Q3: Is/are this/these Constitutive Principle(s) (and the new corresponding definition of necessity) necessary?

It is quite clear that this way of putting things leads to infinite regress. Yet, there is no other way for principle-based accounts (thereby for all typical cases of modal rationalism) to proceed as long as there is no prior unproblematic grasp of necessity. The problem is that this grasp of necessity is required to be not only explicit and unproblematic, but also fundamental – necessity must be taken as primitive if we don't want to set in motion the infinitely regressive necessity – essentiality – necessity – … grounding mechanism.

6. Concluding Remarks

The common thread of the arguments presented here is that they target the main aim of modal rationalism: to provide an account that maintains both that modal truth is mind-independent and, in Peacocke's words, 'fundamentally an a priori matter.' If these arguments are successful, then we have serious reasons to doubt that this task can be carried out. The difficulties of Peacocke's moderate rationalism show that when the grounding a priori knowledge is taken to be implicit (but how could it be otherwise?), it is very unlikely that we will able to characterise our knowledge of metaphysical necessity in an appropriate manner. An adequate modal rationalist account cannot remain programmatic, as it is in Peacocke's work – it has to be fully developed if it aims to fulfill an explanatory role. Otherwise, it is open to doubt and charges of ineffectiveness. But even if such an account is provided, there should be serious concern about its capacity to discriminate between correct and deviant modal knowledge. This has to do with the peculiar nature of the grounding principles, whatever they are taken to be. It is highly doubtful that any principle-based account can garner a large enough consensus, but even if this were to happen, the problem also appears at the second-order level, where the necessity of the account requires a justification of its own.

A somewhat rushed reply would be that the principles are necessary because they are a priori, but anyone that propounds such a solution has missed the most important point. "2+2 = 10" is also a statement about a priori entities, but

in order to deem it necessarily true or false, one has to know first if it is correct or not. This is exactly what we don't know about the principles and the corresponding definitions of modal notions (if we acknowledge that there is an Integration Challenge to be met): are they correct or not? However, as I noted in the previous paragraph, the Principles themselves have a peculiar nature – they don't seem to be able to acquire the same type of consensus that mathematical truths acquire among competent users of mathematics. Now, this raises a very interesting question that is, in my opinion, a fertile challenge for further research. Are the principles of possibility, whatever they are taken to be, a priori and if this is so, what type of a priori knowledge are we talking about? Thinking again about mathematical knowledge, could it be that we operate with modal notions in a similar way, that is, could they project/make explicit certain rules for concept use or certain properties of our concepts?[45] Would modal competence then be a conceptual competence of sorts? The individualised and highly controversial character of some of our modal evaluations raises some concerns about this point of view, but perhaps this concern can be allayed with accepting the fact that different users employ different rules for what appear to be the same concepts. Or, better, modal disagreement may be a consequence of the fact that rules for concept use are not as strict and rigorous as mathematical rules. Naturally, in the case this theoretical option is pursued, it should mark a clear departure from a modal rationalist position; the guideline for assessing the truth of a statement like 'if X is a cat, then X is necessarily an animal' would not be that it correctly tracks some mind-independent truth, but rather that it correctly specifies the limits of use of the concept <cat>. Another option would be to hold that the principles are in fact forms of a posteriori knowledge, perhaps of a more peculiar kind.[46] This is also (and more clearly so) incompatible with a rationalist perspective.

McLeod has argued that ontological realism about modality requires modal rationalism.[47] If this is correct and the problems discussed in this paper truly affect all forms of modal rationalism, then these difficulties are really even more worrying than argued here. I don't wish to pursue this line of reasoning in this paper, but if my inquiry is significant, one thing it clearly suggests is that robust realism about metaphysical modality should be disputed more vigorously in the ongoing epistemological debate regarding modal notions.

---

[45] A view along these lines is developed by Amie Thomasson in "Modal Normativism and the Methods of Metaphysics," *Philosophical Topics* 35 (2007): 135-160 and "Norms and Necessity," *Southern Journal of Philosophy* 51 (2013): 143-160.

[46] See Elder, "An Epistemological Defence of Realism," for a way to argue in favour of this claim.

[47] Cf. McLeod, "Rationalism and Modal Knowledge."

A DAVIDSONIAN RESPONSE TO RADICAL SCEPTICISM

Ju WANG

ABSTRACT: In this paper, I attempt to show how Davidson's anti-sceptical argument can respond to the closureRK-based radical scepticism. My approach will focus on the closureRK principle rather than the possibility that our beliefs could be massively wrong. I first review Davidson's principle of charity and the triangulation argument, and then I extract his theory on content of a belief. According to this theory, content of a belief is determined by its typical cause and other relevant beliefs. With this constraint on content, I argue that doubt must be local. Furthermore, since one cannot rationally believe that one's commitment to the cause of beliefs could be false, our commitment to the denial of a sceptical hypothesis is not a knowledge-apt belief. Therefore, the closureRK principle is not applicable to rational evaluations of this commitment. As a result, the closureRK-based sceptical argument fails while the closureRK principle remains.

KEYWORDS: closureRK principle, radical skepticism, doubt and belief, knowledge-apt belief

0. Introduction

Donald Davidson is a prominent anti-sceptical exemplar in contemporary philosophy. He famously argued that *belief is in its nature veridical*.[1] If the claim were true, one significant presupposition of scepticism would be unattainable. In Davidson's view, scepticism regarding the external world presupposes that our beliefs about the external world could all be false. The idea of massive error in our beliefs can be motivated by considering ordinary cases where our senses present us misleading information. If we grant that a perceptual belief is fallible in particular cases, then it seems that nothing stops us from generalizing such a possibility of error. As far as we cannot exclude the possibility of massive error, scepticism is on the way.[2] Accordingly, Davidson's theory aims to offer compelling reasons as to why such a possibility is unintelligible.

[1] Donald Davidson, "A Coherence Theory of Truth and Knowledge," in his *Subjective, Intersubjective, Objective* (Oxford: Clarendon Press, 2001), 146.

[2] Notice that apart from the idea of massive error, scepticism also needs the closure principle to bring it happen.

The standard formulation of the closure-based sceptical argument consists of the following three claims where S is a subject, Q is the denial of the sceptical hypothesis and P is an everyday proposition:

A. S cannot know that Q.

B. If S knows that P, and S knows that P entails Q, then S knows that Q.

C. S knows that P.

Here is the idea behind this formulation: in order for us to know something, we must be capable of ruling out the possibility of the sceptical hypothesis. If this requirement cannot be met, our epistemic status would be too weak to claim any knowledge. Notice that massive error in beliefs supports claim A implicitly. The reason is simple: in a sceptical scenario, *ex hypothesi*, our beliefs could be massively wrong while we continue believing them to be true. Thus, we cannot know whether the sceptical hypothesis obtains. If Davidson can offer good reasons to refute or dismiss the idea of massive error, then surely we are in a position to tell why claim A is not acceptable in the sceptical argument. Along one line, we can focus on how to interpret Davidson's argument and whether his argument is sound, but I won't pursue this line in this paper. Rather, I attempt to develop Davidson's anti-sceptical argument against claim B. We can respond to the closure-based sceptical argument by rejecting the second claim or what is normally called the closure principle. In order to focus upon a plausible version of B, let's look at a more specific principle formulated by Duncan Pritchard (forthcoming):

The ClosureRK Principle

If S has rationally grounded knowledge that p, and S competently deduces from p that q, thereby forming a belief that q on this basis while retaining her rationally grounded knowledge that p, then S has rationally grounded knowledge that q.[3]

At first sight, this principle is more demanding than the closure principle in that while the closure principle only demands that knowledge be under known entailment, the ClosureRK principle requires that knowledge be based on a rational ground which can preserve across a well-conducted rational operation by a believer. However, the ClosureRK principle is also defensible for this reason: it captures the core idea that knowledge is a cognitive achievement, thus competent deduction, as a paradigm way of a rational process, cannot undermine our rational ground for knowledge. Therefore, what motivates the closure principle motivates

[3] Duncan Pritchard, *Epistemic Angst: Radical Skepticism and the Groundlessness of Our Believing* (Princeton: Princeton University Press, forthcoming in 2015).

the ClosureRK principle as well. Pritchard further argues that what underlies the ClosureRK principle is the idea: the reasons for believing p serve as reasons for believing any logical entailment of p. Pritchard thereby names the idea the universality of reasons thesis.

Contemporary epistemologists generally agree that knowledge entails belief. Specifically, knowledge is responsive to a rational process and belief is knowledge-apt. A knowledge-apt belief is a belief that aims at truth. In what follows, I will limit my discussion to knowledge-apt belief. Given this, a quick point is that if we cannot bear a knowledge-apt-belief attitude towards a proposition, then this proposition is not even in the market for rational knowledge.

I will proceed as follows: In sections 1-2, I review Davidson's theory of radical interpretation, the principle of charity and triangulation. In section 3, I extract Davidson's view on doubt and belief and argue that doubt is in its nature local. What follows in section 4 is an attempt to show how a Davidsonian response to the ClosureRK-based sceptical argument proceeds by denying that the sceptical hypothesis does not obtain is a belief. In the ending section, I give some comments on this solution.

1. Radical Interpretation and Principle of Charity

Davidson appeals to the notion of radical interpretation in order to develop a theory of language, or more broadly, a theory of understanding. What is radical interpretation? Basically, it's a fantasy scenario where one is in a completely new place, say a remote island or an aboriginal tribe. One tries to understand residents by figuring out what they believe and what their utterances mean. However, the interpreter has no prior knowledge of either the speaker's beliefs or the meanings of the speaker's utterances. In such a scenario, what are available to the interpreter are only publicly observable behaviors and utterances, or more specifically, S's holding a sentence true at a certain time and a situation. Thus, the interpreter has to adopt a third-person stance to understand something internal to the speaker. Nonetheless, understanding a speaker seems to be extremely difficult as Davidson remarks, since there is interdependence between a speaker's belief and the meaning of her utterance. Given the interdependence, it follows that the interpreter cannot assign a meaning to a speaker's utterance without specifying the speaker's beliefs, and one cannot identify the speaker's beliefs without figuring out what her utterances mean.

To see this, we need to think what we normally do in everyday life. We express our beliefs by uttering words or sentences. In doing this, we must also be competent to use our idiolect so that we are able to find the appropriate way to

express our beliefs. For instance, if I intend to express my belief that London is a modern city by uttering the sentence "London is a modern city," then I am doing right since I choose the proper sentence to express the corresponding belief. Otherwise, my expression would be improper. Then suppose we are to interpret someone who utters the very sentence, how can we start our interpretation? The problem is we cannot assign the meaning London is a modern city to S's utterance if we are ignorant of what S believes. By the same token, we cannot figure out what S believes unless we already have an idea of what her utterance means. If we plan to assign belief and meaning to a speaker in radical interpretation, the speaker's holding true attitude can be explained by more than one interpretation. Suppose S utters 'X', then I take S to hold her sentence X true and give two interpretations as follows:

(Int1) S believes P	(Int2) S believes Q
X means that P	X means that Q
Hence S holds X true	Hence S holds X true

In (Int1) and (Int2), both interpretations provide an account of why S holds X true based on an ascription of S's utterance and S's belief. However, potential rival interpretations are countless. Thus, it would be very difficult for us even to set off the interpretation for a single utterance. That multiple interpretations are available for one utterance indicates that belief and meaning are co-varied in explaining one's holding true attitude. Without more constraints, we can't determine which interpretation is the optimal.

In order to resolve this problem, Davidson appeals to the principle of charity. Roughly put, the principle of charity requires that the interpreter and the speaker share mostly true and coherent beliefs, by the interpreter's own light. How can this principle help solve the problem of interdependence between belief and meaning? Notice that the interdependence problem arises out of the co-variation of meaning and belief. Therefore, we are supposed to determine meaning while maintaining belief as a constant factor as far as possible. With belief being fixed, we can ascribe meaning to one's utterance, and those ascriptions are not absurd in our own light. For example, let S say "there's a hippopotamus in the refrigerator." Davidson remarks that we prefer interpreting S's utterance as saying that there is an orange in the refrigerator rather than there is a hippopotamus in the refrigerator.[4] The reason is that we find the former interpretation to be more reasonable when S's further account includes descriptions that we normally take

[4] Donald Davidson, "On Saying That," in his *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), 101.

an orange to fit better, rather than a hippopotamus. Notice, here we take our beliefs about orange into consideration when determining the meaning of S's utterance.

Brueckner objects that guaranteeing that the interpreter and the speaker share beliefs as much as possible, as in the orange case, would result in ridiculous interpretations when the speaker and the interpreter live quite different lives, such as when an interpreter from tropical area ascribes a belief to a Siberian resident that he has never seen snow.[5] What's worse, not only is *massive agreement* in beliefs not enough, but also it won't be better if we add that we assign similar beliefs to a speaker by referring to similar *evidence*. The key reason is that my envatted counterpart and I have subjectively indistinguishable perceptual experience, and therefore our evidence for beliefs is on a par. Accordingly, it is required by charity to attribute true and coherent beliefs to her, but this move is at odds with my intuition that a BIV cannot have true beliefs about her environment. Moreover, unless we as interpreters have generally true beliefs, our attributed beliefs to others could not be true. In other words, it makes no sense to assign true beliefs if interpreters are massively wrong. Thus, there is something wanting in defense of the principle of charity, and Davidson is also aware of the weakness. Later on, Davidson puts forward the omniscient argument, but it's not what we should pursue here because he renders it as a useless argument on the one hand, and there are fatal problems with this argument on the other hand.[6] Therefore, I will proceed to Davidson's triangulation argument which takes over the role of omniscient argument as articulating and defending the principle of charity.

2. Triangulation

What is lacking in the previous formulation of charity is that, only massive agreement and similar evidence cannot make beliefs generally true. A crucial problem is if we can't determine the cause of beliefs, it is still possible that our beliefs are mistaken in the main since our beliefs may be caused by the evil demon rather than an objective world. Consequently, Davidson has to settle the issue how a belief gets its content. In "Rational Animals," Davidson suggests a triangulation model of radical interpretation:

---

[5] Anthony Brueckner, "Charity and Skepticism," *Pacific Philosophical Quarterly* 67 (1986): 264-68.

[6] Donald Davidson, "Reply to A. C. Genova," in *The Philosophy of Donald Davidson*, ed. Lewis Hahn (Chicago: Open Court, 1999), 192-194.

If I were bolted to the earth, I would have no way of determining the distance from me of many objects. I would only know they were on some line drawn from me towards them. I might interact successfully with objects, but I could have no way of giving content to the question where they were. Not being bolted down, I am free to triangulate, one that requires two creatures. Each interacts with an object, but what gives each the concept of the way things are objectively is the base line formed between the creatures by language. The fact that they share a concept of truth alone makes sense of the claim that they have beliefs, that they are able to assign objects a place in the public world.[7]

A triangulation, just like a triangle in geometry, involves three points and three lines. It consists of an interpreter I, a speaker S, an object O, and lines connecting I, S and O respectively. I-O and S-O are both causal relations between a subject and the external world, while S-I forms a social network between two subjects. In triangulation, both subjects respond to the object located in a shared world and react to each other's reactions. Why appeal to such an argument? I suppose Davidson is in need of a new argument to justify the principle of charity and to establish that belief is in its nature veridical. How can he achieve this task via triangulation? Davidson urges that,

> we must, in the plainest and methodologically most basic cases, *take the objects of a belief to be the causes of that belief*.[8]

However, in determining the cause of a belief, we have to face the ambiguity of the concept of cause. The ambiguity is two-fold. On the one hand, one cannot locate the object that causes one's belief. The cause of one's belief may be distal or proximal, and it may lie at any point along the I-O line; on the other hand, one cannot determine what aspects of the object typically cause one's belief. In triangulation, what is essential is that when two causal lines converge, each can locate the cause of their belief. The reason is that what causes my belief may stand at any point along the I-O line, and so does whatever causes your belief. Nonetheless, what causes your belief can't be something in my mind or stimuli in by brain (i.e., proximal), therefore the cause of my belief must be somewhere in the external environment where our causal lines meet. Is the convergence enough to specify the cause of belief? The answer is no. Davidson needs the I-S line to identify the cause of belief as well. To see this, recall the orange in the refrigerator case. Even interpreter and speaker can locate where the cause of their beliefs is,

[7] Donald Davidson, "Rational Animals," in *Subjective, Intersubjective, Objective* (Oxford: Clarendon Press, 2001), 105.

[8] Davidson, "A Coherence Theory," 151.

they have a disagreement about what the object is. Only through further descriptions, they can settle whether the cause is an orange or a hippopotamus. Those further descriptions are not only further information that can disambiguate the cause of belief, but also a requirement for a rational belief. The role of I-S line is significantly vital in two aspects.

Firstly, it provides the space of error. Since only in communication, both interpreter and speaker are in a position to identify the cause of their beliefs. When each alone attributes the cause of beliefs or when they assign beliefs to each other without communication, they sometimes make mistakes. However, it is implausible to correct this mistake by oneself since one cannot distinguish what is true from what is false. Communication provides a space in which one could realize that the truth of one's belief is determined by an objective world which one shares with others.

Secondly, it seems that what is at issue here is language communication. It is true that we can identify the cause of a belief only in communication, but communication is not only a matter of language, it's also a matter of thought. In order to specify an orange as the cause of a belief, the interpreter considers several sentences of the speaker, not only the first claim that "there's a hippopotamus in the refrigerator." Those sentences help to determine whether the speaker is talking about an orange or a hippopotamus. Meanwhile, given that sentences express thoughts, we can infer that if the speaker's belief is about an orange, a set of further beliefs should also be endorsed by the speaker.

Obviously, Davidson is here invoking semantic holism. Semantic holism is the thesis that meaning of a sentence is determined by other related sentences in the language. Even though it is impossible to draw a clear line where related sentences end, I suppose the core idea behind it is still plausible. We can motivate semantic holism by considering the indeterminacy of interpretation.[9] In developing a theory of meaning for a language, one must deal with the totality of utterances in that language. However, given a certain finite amount of evidence, we can have more than one theory of interpretation that suffices to account for the evidence. Different interpretations can differ in a particular assignment of meaning and belief, but would be equally plausible in explaining the total evidence. It is the whole sentences and beliefs that have equally satisfying explanation to one's action regardless of differences in minor places. What follows from the indeterminacy of interpretation is the holistic nature of meaning and belief rather than a practical or cognitive limitation in interpretation. Another

[9] Jeff Malpas, "Donald Davidson," *The Stanford Encyclopedia of Philosophy*, Summer 2014 Edition, http://plato.stanford.edu/archives/sum2014/entries/davidson/.

reason to support semantic holism is to appeal to the inferential property between meaning and belief. For instance, 'orange' is inferentially connected to 'fruit,' and further inferentially connected to 'non-meat'; similarly, my belief that this is an orange is inferentially connected to my belief that this is a fruit and that this is not meat. The holistic structure of meaning and belief not only serve as an explanatory tool in accounting for one's mental content, but also constitute a normative regulation on any attribution of meaning and belief. Therefore, to entertain a singular proposition is not possible in Davidson's view; at least the rationality of thought and language can only be appreciated in a holistic structure.

To summarize, in the triangulation model, the content of a belief is determined by two factors. The first is the cause of the belief, which is indicated by I-O and S-O lines; the second is other relevant beliefs in the belief system, which is also a social trait embodied in the S-O line. So far, Davidson may have plenty of resources to argue that belief is in its nature veridical, and surely one can proceed to argue against the possibility of massive error. However, I will bear in mind his view on the content of belief and then turn to his comment on doubt and belief.

3. Belief, Doubt and Content

When it comes to beliefs, Davidson emphasizes that a belief must have a determinate content. What determines a belief's content then? According to the previous discussion, we can specify the content of a belief by relating it to a world of further beliefs or its typical cause. See the following remark of Davidson:

> Not only does each belief require a world of further beliefs to give it content and identity, but every other propositional attitude depends for its particularity on a similar world of beliefs.[10]

By relating a belief to other beliefs, we can determine the very content of the belief in question. This claim is also supported by the semantic holism endorsed by Davidson. And since other propositional attitudes, like doubt, hope and wishful thinking, etc. depend on beliefs, we can only proceed to doubt by having beliefs in the first place.

Davidson thus gives priority to belief among different propositional attitudes. Why is this plausible? To see this, let's have a look at the following example. When S doubts that there is a tree in the garden, she must have something in mind to support her doubt. These things in mind are beliefs she actually holds, like that a tree is an object with branches and trunk and that a

[10] Davidson, "Rational Animals," 99.

garden is a space where plants are cultivated and displayed. Without those further beliefs in play, her doubt regarding the proposition would be void of content and thereby nonsense. The same holds for other propositional attitudes like hope that there is a tree in the garden, or fear that there is a tree in the garden. The core idea is that, in order to have a propositional attitude towards a proposition, or having a thought, it is necessary to have a network of related beliefs which serve to make the propositional attitude in question intelligible and contentful. Therefore, the starting point of any propositional attitude is belief, and so is doubt.

In terms of doubt, Davidson proposes a constraint:

> It is only after belief has a content that it can be doubted. Only in the context of a system tied to the world can a doubt be formulated.[11]

Davidson's claim here can be split up into two aspects. One is how to have a doubt; the other is how to formulate a doubt if we had it. However, on closer inspection, I suspect Davidson has only one thing in mind. It is obvious that the logic of doubt, according to Davidson, presupposes belief with a content. That is to say; it is not possible to doubt that p unless p has a content, and it is only by having a prior set of relevant contentful beliefs that one can give this doubt a content. This claim is a natural consequence of the previous discussion on the order of propositional attitude, which says that belief is prior to doubt.

Suppose then we form a genuine doubt that p by having a prior set of relevant contentful beliefs, why do we have to appeal to a context of system tied to the world in order to formulate the doubt? Won't it be enough for us to have a set of beliefs, worldy or otherwise, that can help us to determine the content of the proposition being doubted? Here, we need to return to the question, in virtue of what the content of a belief is determined. I have made it clear that both a set of further beliefs and the typical cause of a belief contribute to determining the content of a belief. However, the content-determining role of the typical cause is different from that played by the further relevant beliefs. In order to guarantee that our beliefs are objective, coherence between beliefs does not suffice. What is essential is that those beliefs should be about the external world, i.e., the origin of the content of beliefs is the external world. Thus, if we intend to formulate a doubt that p, we use our language to express this thought. Given that Davidson argues that thought and language are interdependent, it follows that the same content entertained by thought can be expressed by language. To sum up, belief and sentences owe its original content to the external world.

[11] Donald Davidson, "Reply to Barry Stroud," in *The Philosophy of Donald Davidson*, 165.

In what follows, I will explore the consequence of Davidson's views about doubt and belief. It seems that what is entailed by Davidson's views in this respect is that doubt can't be wholesale. Here is the reason. In order for S to have a doubt that p, she must identify the content of her doubt that p. However, S can do this only by relating p to other relevant contentful beliefs which she holds; that is to say, S is required to believe for her doubt to take place. Also, since other relevant beliefs are connected to an external world, a doubt is therefore about an external world as well. Here, what S believes is fundamental to what S doubts, so S's doubt presupposes S's belief. The idea that there is a wholesale doubt is simply unintelligible, for a wholesale doubt puts all one's beliefs into question and therefore we can't find any leftover beliefs to determine the content of those propositions being doubted. The result is that the idea of a wholesale doubt simply can't get off the ground. This is so not because of a constraint on our cognitive capability, but rather because the essence of doubt excludes a wholesale doubt. A wholesale doubt is impossible, and thereby doubt must be in its nature local.

4. Consider the ClosureRK Principle Again

Recall the ClosureRK principle, and let's take "There's a rabbit in front of me" as p and "I am not a BIV" as q; then an instance of ClosureRK would be as follows:

> *If S has rational grounded knowledge that there is a rabbit in front of me and S competently deduces belief that I am not a BIV via knowledge that there is a rabbit in front of me, then S has rational grounded knowledge that I am not a BIV.*

Given what I have discussed, I doubt whether this instance is plausible. It fails not because ClosureRK fails, but rather that the instantiation of q is rather dubitable in this case. Specifically, I am wondering whether S could have a belief attitude towards the proposition that I am not a BIV. If not, this proposition would lie beyond the realm of rational knowledge. However, it is still not clear why that I am not a BIV doesn't count as a belief.

What is necessary to have a thought? Or more specifically, what does it take to have a belief? Apart from what we have discussed earlier, Davidson maintains that:

> In order to have a belief, it is necessary to have the concept of belief.[12]

What is the concept of belief? One may suggest that the concept of a belief is the meaning of the concept 'belief.' In this sense, having the concept of belief is

[12] Davidson, "Rational Animals," 104.

a precondition to use the word 'belief' in a sentence, which is similar to understanding the meaning of 'desk' before using the word in a sentence. However, this is not what Davidson has in mind. Davidson won't deny that there is a strong relationship between language and thought, so that in order to have a thought such as a belief, S must be able to express that belief in language. But the core idea of what is the concept of a belief is something other than this.

We can articulate the concept of a belief, as Davidson later suggests, by referring to the concept of objective truth. Objective truth requires that the truth of a belief is not determined by any of the subject's mental states. Rather, a belief owes its truth partly to the external world. Accordingly, the truth of my belief is not determined merely by my mental states, and it is quite natural to assess my belief in part by reference to how the world is. If a belief is capable of being true or false, correct of incorrect, then to have a belief that p is different from it is true that P. After all, it is always possible that what I believe is not the case. How can we cash out the possibility that my belief could be wrong? Davidson emphasizes that linguistic communication could show command of the contrast between what is believed and what is the case. Further, I contend that the possibility of a thought's being true or false is best illustrated by (i) making sense of the possibility by appealing to relevant beliefs that S holds and (ii) [….] from S's perspective.

Why is this move acceptable? The claim (i) is based on Davidson's idea just mentioned. In linguistic communication, when the difference between what is believed and what is the case is revealed, S who finds her prior belief was wrong must be possible to make this wrong belief intelligible, otherwise there cannot be such thing as a recognizable error. This error is intelligible just because S's beliefs in stock can account for such a mistake, and can accommodate a revised new belief. As to claim (ii), why making sense of the possibility has to be from S's perspective? The simple reason is: thought is always someone's thought, thus when attributing thought to someone S, we ought to consider if this thought can be held from S's own light. In other words, we should take into consideration whether S has the network of beliefs to maintain the belief in question. It is for this reason that Davidson remarks:

> I reserve the word 'concept' for cases where it makes clear sense to speak of a mistake, a mistake not only as seen from an intelligent observer's point of view, but as seen from the creature's point of view.[13]

Here is a case in point. If I believe that the Lake District is in Scotland, I should be aware that the belief could be wrong. What marks my awareness then?

[13] Donald Davidson, *Problems of Rationality* (Oxford: Clarendon Press, 2004), 9.

Is it the claim that my belief could be wrong? Or is it something further that could make the supposed wrong belief intelligible? Awareness of this possibility of being wrong does not bring us any closer to the intelligibility of being wrong, we have to establish a room for the belief to be wrong. How is this room to be built? From my perspective and what Davidson has argued, the room of false belief rests on many other true beliefs. To realize that this belief could be wrong does not require us to find more evidence to support this claim. What is at stake is that it must be possible to make sense of when the belief is regarded as false. My example continues: suppose my belief that the Lake District is in Scotland is wrong (it is actually wrong), its being false would be intelligible for the reason that I had misidentified Lomond National Park, which is instead located in Scotland, as the Lake District, or the reason that I was ignorant of the boundary between Scotland and England, which caused my failure to recognize that the Lake District is a part of England. In this case, my false belief remains possible because it can be explained by other beliefs I still hold, such as that I believe that I had misidentified Lomond National Park as the Lake District, or that I believe that I was unsure where the boundary between Scotland and England lies. The point is that a certain belief's being wrong does not lead to a breakdown in my belief system so that I still have beliefs to which I appeal to account for this possibility.

Here, I need to clarify what the breakdown means. When we find a prior belief to be wrong, it is natural for us to explain this situation by appealing to some undoubted beliefs, or a step further, to adjust to this change by changing some beliefs consequently. Changing the epistemic status of a belief, i.e. being true, being false and doubtful etc., may result in a chain of changes in our belief system, and the degree of change is determined by its status in the belief system. What I contend here may sound like a Quinean picture of knowledge,[14] but I shall limit my discussion to the content of beliefs only, i.e. what may cause a drastic change in the content of beliefs in the belief system. I would say nothing about adopting or abandoning beliefs for practical reasons or moral reasons.[15] To

[14] According to Quine, our whole system of scientific claims, or our "*web of belief*," faces the tribunal of experience as a unity. When encountered with recalcitrant experience, or experience that does not fit our expectations, we could change observation beliefs, meanings of beliefs or even logical laws to accommodate this change. Essentially, nothing in our belief system is immune to revision. However, logical laws or fundamental hypotheses of a theory are so important that any change in them would result in massive changes in our belief system. Restrained by the maxim of minimum mutilation, we tend to avoid drastic changes and make adjustments in other places.

[15] For instance, a doctor may, for the patient's benefit, tell her something that isn't true. In this case, a false belief is adopted for practical reasons. Likewise, we tend to think that cheating is

summarize, what I mean by breakdown of a belief system is a drastic change in the content of beliefs in the system. The crux is what can cause such a breakdown? In my view, the answer can be easily found in the origin of content of beliefs. Since a belief gets its content from other related beliefs and external causal object, what causes a breakdown in a system must be fundamental commitments in these two aspects, i.e., those about the relations between beliefs and those regarding the external world. The former are logical truths, and the latter are our commitments to the external world or its sceptical counterpart. There is no wonder that a change in logical truth such as declining the law of identity would have a devastating consequence in our belief system. But it is not equally clear what if a change occurs in our commitments to the external world, or to the sceptical scenario for a sceptic. Let's see the following pair:

I am a BIV = the sceptical hypothesis does obtain

I am not a BIV =the sceptical hypothesis does not obtain

From the non-sceptical perspective, we are committed to the sceptical hypothesis not obtaining, i.e., that I am not a BIV. What comes next is a supposed change in this commitment. Our commitment changes to I am a BIV, or that the sceptical hypothesis does obtain. What follows this change? Apparently, our ordinary beliefs regarding the allegedly external object are all changed since the causes of beliefs are only stimuli generated by a supercomputer. It seems that we can still imagine what it is to be a BIV, but there is surely something wrong with this. In normal circumstances, when we talk about "there is a tree in front of me," this belief is caused by and therefore about a tree standing in front of the speaker/believer; while in BIV case, we find it plausible that the same sentence would be caused by and thereby about a series of stimuli regardless of the same subjective perceptual experience. However, what makes our claim about the change of content plausible? It is not that someone had been or is now in the exact sceptical situation, but that some cases we encountered in normal circumstances, like dreams and hallucinations, underlie the plausibility of such claim locally. Nonetheless, the local plausibility does not easily extend to universal, so the sceptical hypotheses usually rest on a thin sense that a sceptical scenario is not a logic contradiction thus it is metaphysically possibility. However, if I have been always a BIV from birth, chances are that I could never see that my beliefs are

---

morally wrong. Even though cheating can bring benefits to an individual at times, given certain moral considerations such as that we should be righteous and we should treat others fairly, we have sufficient moral reasons to abandon the belief that cheating is a good behaviour. The point is that we change beliefs for non-epistemic reasons in everyday life.

about stimuli generated by a supercomputer. Even though the supercomputer can bestow me any stimuli needed, my belief can never get content about the computer. What's more, I couldn't have an idea of what the objective world is. Since in order to get the content from the objective world, BIV must be isolated from the supercomputer, otherwise the belief would still be about stimuli rather than worldly objects. It is obvious that by changing our fundamental commitment to the cause of our beliefs, we are committed to change the content of our ordinary beliefs as well. What is crucial now is that, from my perspective, all my previous beliefs would lose their content since they couldn't be about the external world as I earlier presumed. However, I cannot make sense any of the false belief because I have no any contentful beliefs left to which I appeal, thus this possibility of its being false, or the possibility that I am a BIV cannot be genuinely appreciated by me. If I cannot appreciate the possibility of being false, then the original proposition that I am not a BIV does not qualify as a belief at all. It does not count as a belief because no one can appreciate this proposition's being false with determinate content.

Interestingly, the case would be likewise if we start from a sceptical hypothesis. If we start with the commitment that I am a BIV, then supposing the possibility that I am not a BIV would also change the content of all beliefs already formed. All my beliefs are about computer stimuli, and none of them is caused by the external world. If I were to form any belief about the external world due to the change of my commitment, I would have to sweep all my previous beliefs and form new basic beliefs from the scratch. Since no any contentful belief remains to help me identify the content of my belief that I am not a BIV, this initial commitment fails to be a belief either. Thus, it seems that no matter what we are committed to at first, whether I am a BIV or not, as far as this commitment leads to a massive change of the content of beliefs, it cannot be a knowledge-apt belief at all.

To sum up, in order to have a knowledge-apt belief, we must have the concept of knowledge-apt belief. We can command the concept of knowledge-apt belief by having the concept of objective truth. Objective truth requires that we are aware of and are capable of appreciating the possibility of a belief's being true or false. And this possibility cannot be appreciated unless we have some related contentful beliefs to identity the content of the very belief. Crucially, a change in our fundamental commitment to the cause of beliefs would result in a total change of content in all beliefs, which leaves us no contentful belief at all to make this possibility intelligible. Therefore, it is impossible to appreciate the possibility of our fundamental commitment's being false from our own perspective. It is now

clear that our commitment does not meet the requirement of having a belief. And it follows that the sceptical hypothesis does not obtain is not a knowledge-apt belief. Thus, the ClosureRK principle is not applicable to the evaluation of the sceptical hypothesis.

5. Conclusion

I have argued that our commitment to the denial of the sceptical hypothesis is not a rational belief; therefore it cannot be applied to the ClosureRK principle. But it is still not clear why sceptics and non-sceptical epistemologists take it as a belief? The therapeutic answer from Davidson is that they both misconceive how a belief can maintain its content while changing its content determining conditions. For sceptics, they wrongly contend that the content of beliefs can be the same while the cause of beliefs changed dramatically and that a universal doubt is a contentful doubt. Therefore, they hold a belief-like attitude towards the proposition that the sceptical hypothesis does not obtain via having other genuine contentful everyday belief. Meanwhile, it is tempting to think that our increased success of attributing ordinary beliefs to others, may serve as a reason to justify that the sceptical hypothesis does not obtain is very likely to be true. However, I am afraid that this idea also goes against Davidson's claim that it is absurd to look for a justifying ground for the totality of beliefs. Our commitment to the cause of beliefs serve as a ground to attribute content to all beliefs, but there is no need to justify this ground since it is in nature outside the reach of any rational beliefs and therefore non-justifiable.

Davidson arrives at the conclusion that doubt is local, and this conclusion echoes Wittgenstein's comment on the structure of reasons. Thus, I would further explore the details between these two diagnoses. Regarding the nature of doubt, Wittgenstein claims:

> [...] the questions that we raise and our doubts depend upon the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn. That is to say, it belongs to the logic of our scientific investigations that certain things are indeed not doubted. But it isn't that the situation is like this: We just can't investigate everything, and for that reason we are forced to rest content with assumption. If I want the door to turn, the hinges must stay put.[16]

In this quotation, Wittgenstein maintains that it is only possible to doubt when some propositions are undoubted. Wittgenstein advances his argument from considering the certainty of propositions. Propositions have different degree of

[16] Ludwig Wittgenstein, *On Certainty*, eds. G. E. M. Anscombe and G. H. von Wright, trans. D. Paul an G. E. M. Anscombe (Oxford: Blackwell, 1969), §§341-3.

certainties, some of which we are pretty sure while others are not. Thus, propositions with the highest certainty can provide rational support for those which have less certainty, but it is because of the highest certainty that those propositions are groundless. They are groundless because no propositions are more certain to provide rational support for them. What follows is that there can't be rational doubt regarding those hinge propositions on which any rational evaluations take place. In a nutshell, doubt must be local in its nature.

Even though Davidson and Wittgenstein arrive at the same conclusion, reasons for the conclusion are different. Davidson draws this conclusion from a consideration on semantic/mental content, i.e., what makes a doubt with a determinate content? In contrast, Wittgenstein focusses upon the epistemic ground of doubt, i.e., what makes a doubt epistemically rational? It is natural to ask whether a contentful doubt is epistemically grounded. Suppose I doubt that this is an orange, I may have plenty of content-determining beliefs that an orange is a round shape fruit, juicy and tastes sweet and sour. Those beliefs can give my doubt a content about an orange rather than a rock. But it is likely that none of these beliefs serves as an epistemic ground to support my doubt. Since to rationally doubt this is an orange, I need to have some evidence like an experienced and reliable farmer told me that it's not an orange. Conversely, an epistemically grounded doubt may not come with a determinate content. If I was informed by an experienced and reliable farmer that this is not an orange, then I can rationally doubt that it is an orange. However, if I don't have belief that an orange is a round shape fruit, juicy and tastes sweet and sour, I may continue believing that this is a round shape fruit, juicy and tastes sweet and sour while maintaining the doubt that this is an orange, which is very implausible. In this case, people would say my doubt is not about an orange since I believe other propositions that fit perfectly with the orange in front of me. One may find it more plausible to argue that the farmer's information is not reliable, or that 'orange' in the farmer's language is not the same as in our language. My point is, if we don't have relevant beliefs to determine the content of a doubt, it would be impossible to appreciate any epistemic support for your doubt. A moral we can draw is that the content requirement of a doubt doesn't coincide with its epistemic support. However, a further question is, in order to rationally doubt something, do we have to have both content and epistemic support in place? If the answer is yes, then Wittgenstein and Davidson could go hand in hand in their respective responses to the closure-based sceptical argument.

I suppose here is a reason to answer positively. In order to have a rational doubt, it is not only a matter of having a contentful doubt, but also having a

rationally supported doubt. Content of the doubt determines what the subject of the doubt is, and epistemic support for the doubt determines whether the doubt about something is well-motivated. If I aim to have a rational doubt on my friend Weighty's weight, I need to know first what weight is and what Weighty's weight is to be understood. Besides, I also need to know what ground I have in order to doubt Weighty's weight. The first requirement can guarantee that my propositional attitude is really about Weighty's weight rather than Weighty's height or Heighty's weight, while the second requirement can form this propositional attitude as a doubt, rather than believing or wishful thinking. A doubt without determinate content cannot be distinguished from other doubt, and a doubt without a rational support cannot be a doubting attitude at all. Accordingly, those two requirements are not only compatible, but also mutually supported. In this sense, doubt is essentially local both in terms of its content and its epistemic support. Indeed, Wittgenstein and Davidson could have a unified diagnosis of the sceptical illusion that a wholesale doubt is attainable.

It is now clear that while the ClosureRK principle is plausible, it does not licence the sceptical conclusion. What is pivotal is that this principle is only applicable to (knowledge-apt) beliefs. Unfortunately, our commitment to the denial of the sceptical hypothesis is not a (knowledge-apt) belief. From a Davidsonian perspective, a (knowledge-apt) belief essentially involves other content-giving beliefs. Commitment to the denial of the sceptical hypothesis is so fundamental that, when being doubted, any contentful beliefs would be discarded therefore. Hence, it will be impossible to take it as a rational, contentful and (knowledge-apt) belief. As a result, we can retain the ClosureRK principle while evading the sceptical challenge apparently posed by this principle.[17]

[17] I am grateful to Duncan Pritchard and Allan Hazlett for helpful discussions. My research is funded by China Scholarship Council and University of Edinburgh.

DISCUSSION NOTES/DEBATE

DISAGREEMENT AND PHILOSOPHICAL PROGRESS

Brent ABLES

ABSTRACT: In "Belief in the Face of Controversy," Hilary Kornblith argues for a radical form of epistemic modesty: given that there has been no demonstrable cumulative progress in the history of philosophy – as there has been in formal logic, math, and science – Kornblith concludes that philosophers do not have the epistemic credibility to be trusted as authorities on the questions they attempt to answer. After reconstructing Kornblith's position, I will suggest that it requires us to adopt a different conception of philosophy's epistemic value. First, I will argue that 'progress' has a different meaning in logic, science and philosophy, and that to judge one of these disciplines by the standards appropriate to one of the others obscures the unique epistemic functions of all. Second, I will argue that philosophy is epistemically unique in that it is a non-relativistic but historically determined excavation of foundations. Finally, drawing on Frank Herbert's *Dune,* I will suggest that Kornblith leaves us with a choice between two epistemic ideals: the hyper-logical 'Mentat,' or the historically informed 'pre-born.'

KEYWORDS: disagreement, epistemic modesty, epistemic authority, Hilary Kornblith

1. Introduction: The Mentat Ideal

In Frank Herbert's popular series of science fiction novels, *The Dune Chronicles,* humankind has done away with complex machinery and conditioned itself to fill the void. Instead of computers, there are 'Mentats.' A Mentat is a human being with extraordinary logical capacities, able to solve extremely complex problems in a short time and make predictions with a very high degree of empirical accuracy. They receive information as input, and their output is truth. Mentats are still human, however; they marry, exhibit loyalty to their authorities, and have complex emotions. But if they begin to get carried away by those emotions, they need only remind themselves to "function as a Mentat" and their dominant rational side reasserts itself.[1]

We can probably agree on two facts about Mentats. The first is that they seem like an excellent ideal for the philosopher to uphold. To be able to stop what one is doing, set aside emotions and passions, and function as a Mentat is surely a skill that any honest seeker of truth would want to cultivate. In contemporary

[1] Frank Herbert, *Dune* (New York: Penguin, 1965), 17.

analytic epistemology, the paradigm of the epistemic agent is often something very like a Mentat. It is supposed that we receive evidence as input, process that evidence in accordance with the best methods and standards that we have available to us, and form true beliefs as output. Even if it is acknowledged that real-world circumstances might interfere with this process, it is taken by epistemologists of disagreement to be a suitable norm.

The second fact is that Mentats belong in science fiction. As much as we would like to be able to cordon off error and the illusions of what William James called our passional nature, we know from experience that willpower and practice cannot reliably accomplish this. Look at the methodology of contemporary science, which demands of any individual researcher a full explanation of the procedures and materials she employs so that her experimental results can be replicated and verified by other scientists. Or take the academic process of peer review, which would hardly be necessary if we could all trust our qualified peers to reach proper conclusions given the same data that we have. These kinds of institutional checks are necessary because academics and scientists can be dishonest, but more fundamentally because we recognize that in the collective human search for truth, even the brightest and most honest individuals among us can – and often do – get things wrong. The Mentat is an artifact of imaginative fiction because the ideal is simply too good to be true.

The epistemology of disagreement takes this constitutive imperfection of the rational animal as its starting point. If we were all capable of functioning as Mentats on cue, then substantive philosophical disagreements – indeed, any disagreements whatsoever that go beyond expressions of personal preference – would be systematically resolvable in short order. All the participants would have to do is feed each other the evidence and arguments available, and one of the parties would recognize their error. But this isn't, of course, what happens in the actual practice of philosophy. What happens is that we ask ourselves and each other a varied but limited number of fundamental questions, draw out problems and further questions from that basis, and then argue interminably about those questions and problems at conferences, in journals and private conversations, and even in our heads. As a matter of fact, to paraphrase James, we find ourselves disagreeing.

What does this disjunct between the Mentat model and the messy human reality mean for philosophy? What conclusions ought we to draw about a practice whose results seem so apparently at odds with its ideals? Can these conclusions tell us what to do in the case of substantive disagreement with our colleagues? What is the philosophical significance of such disagreement? In "Belief in the Face of

Controversy," Hilary Kornblith confronts these questions head-on and ends up with some rather pessimistic conclusions about the epistemic authority of philosophy.[2] My goal in this essay shall be to evaluate these conclusions and to determine, not just what philosophy can tell us about disagreement, but what disagreement can tell us about philosophy.

2. Kornblith: The Question of Progress

Kornblith begins his essay by delimiting the scope of the epistemological problem of disagreement. This reduction is necessary since, at first glance, disagreement is as ubiquitous as belief itself: from politics to religion to simple matters of geographical fact, it seems there is almost no topic about which people do not disagree. Not all disagreements are epistemologically significant, however. In some cases, I may have evidence that others do not have and know (or be reasonably certain) that if they had the same evidence, they would believe as I do. In other situations, I may have good reasons to doubt someone's judgment even if we have the same evidence. If I disagree about a simple addition problem with a child learning arithmetic, clearly it would not be proper for me to defer to the child's belief – no matter how certain she may be. Thus, in cases where I have good reasons to doubt my peer's judgment or their access to relevant evidence, disagreement poses no problem. Conviction itself is no guarantee of truth, or even a reliable indication of it.

The epistemologically significant cases of disagreement are those concerning what Kornblith, following Gary Gutting, calls epistemic peers. If you and I are equally intelligent and well-educated, have access to and familiarity with the same evidence, and are both sincere about our beliefs, then we are epistemic peers. These three requirements are jointly necessary, and none alone is sufficient. If we add the proviso that the same interlocutor might be an epistemic peer with regard to one intellectual domain and not another, then we have the idealized interpersonal scenario presupposed by epistemologists of disagreement: a dialogue between humans doing their best to function as Mentats.

To be sure, not all writers who make use of the idea of epistemic peerhood use it the same way. One of the more interesting complications of the idea is raised by Adam Elga.[3] In addition to the criteria already discussed, Elga adds that epistemic peers must have broadly similar beliefs with regard to the topic under dispute. If you and I are both intelligent and generally well-educated, and we find

[2] Hilary Kornblith, "Belief in the Face of Controversy," in *Disagreement,* ed. Richard Feldman and Ted A. Warfield (New York: Oxford University Press, 2010), 29-52.

[3] See Adam Elga, "Reflection and Disagreement," *Nous* 41, 3 (2007): 478-502.

ourselves disagreeing about whether the animal that runs in front of us is a squirrel or a chipmunk, then I can disagree you with in this case and still consider you my epistemic peer. It may be that neither of us is especially well-qualified to discuss the physiology of small furry animals, but we nevertheless both have certain beliefs about other facts regarding chipmunks and squirrels: that they have tails and fur, they can climb trees, etc. The question could be resolved simply by reference to an illustrated encyclopedia.

If, on the other hand, you are an idealist in the tradition of Berkley and I a strict materialist, then our disagreement about what we see is going to be less amenable to resolution. When you remark on the cute notion in God's mind that illusorily appeared to dart by, I might respond by saying that all I saw was a bundle of furry flesh traversing a plot of space-time. According to Elga, this disagreement is epistemically benign: because we are bringing such radically different sets of beliefs to the table, then our ground for considering each other as epistemic peers dissolves under us. Once we set aside all our differences pertaining to the question at hand, there remains no fact of the matter that would allow me to judge my faculty of judgment against yours. And without shared standards for comparison, the concept of epistemic peerhood is meaningless.

Kornblith rejects Elga's additional criterion for epistemic peerhood. He argues that even in cases where beliefs appear to diverge so radically, there are nonetheless still common standards that both parties can appeal to. You and I may disagree about the metaphysical constitution of the critter in front of us, but we both agree that it is not a duck or a lacrosse stick; we both agree that it appeared to run from right to left; we both agree that it appeared to have four legs. Thus, I still have shared grounds for considering you my epistemic peer.

Now, Elga might be willing to concede the point in this case. But surely, he might respond, there are other plausible cases where our relevant clusters of belief really do differ radically and irreconcilably. Kornblith's rebuttal attempts to draw a line between degrees of belief-similarity, and where there are degrees there are exceptions to any fixed rule. But a stronger rebuttal to Elga would be to insist that epistemic peerhood doesn't have anything to do with the *content* of beliefs at all. If you were not already my epistemic peer prior to our sighting the furry animal, then surely you wouldn't become one upon agreeing with me about what we have seen; conversely, if you were already my peer, then I ought not revoke this status when I discover that where I saw a squirrel running you saw a lacrosse stick being thrown. Certainly *some* response to the discrepancy is warranted, but as Kornblith suggests in the case of disagreements regarding perception, I have no more reason *prima facie* to think it is your error than mine. Generalizing the point, epistemic

peerhood is a matter of our trusting the judgmental faculties of another, not endorsing their particular conclusions. It is a sign of methodological competence, not empirical adequacy.

For now, however, let us accept Kornblith's understanding of epistemic peerhood. The question of his essay is: what ought we as peers to do when we genuinely disagree with each other? Before examining Kornblith's answer, we note two important facts about the question itself. First, it is not a descriptive question but a normative one. As much as philosophers like to appeal to thought experiments and counterexamples to grease the axles of intuition, no appeals to juries or dinner bill calculations can answer the question on their own. Values must enter the discussion if it is to go anywhere at all. The second fact is that we can't appeal to some kind of latent but abstract error in calculation to come down on one side or the other. In other words, I can't answer the question by saying that if I am right then I should hold my ground, but if I am wrong then I should suspend judgment. No doubt this is true, but it evades the terms of the question inasmuch as both parties initially take their views to be right. The symmetry of epistemic peerhood must not be broken.

How, then, does Kornblith advise us to act in the case of genuine disagreement with epistemic peers? Let's take the first form of the example discussed above. If I see a chipmunk running across the ground and you see a squirrel, clearly one of us cannot be right. The animal can't be both chipmunk and squirrel, so barring other relevant possibilities (the animal might be a groundhog) the law of contradiction demands that we come down on one side or another. Nevertheless, neither of us can support our belief with any immediate evidence beyond the simple attestation of perception. Thus, Kornblith concludes, the rational thing for us to do is to suspend judgment until we can consult the encyclopedia. Similarly, if you and I add up a dinner bill and come up with two different sums, one of us must be wrong; the logical thing to do in this case would be to suspend immediate judgment and recalculate the bill separately.

What about philosophical disagreement? The widespread presence of disagreement in philosophy is evident from a look at any elementary ethics or metaphysics textbook. But what differentiates philosophy from simple perceptual or mathematical cases, according to Kornblith, is that disagreements in perception or math will or can be easily resolvable. All we have to do is look closer, or redo the calculations. But when philosophers disagree, they usually do so after having examined and considered all the relevant arguments. If they are truly peers, an externalist can't simply tell the internalist to look closer at the arguments to be convinced, nor can she appeal to some impartial third party as an authority (since

she and her peer are as authoritative as anyone else on the subject at hand). There may be no impartial way to resolve the dispute, even if we accept that both sides cannot be right. Granted, then, that philosophical disagreements differ importantly from more mundane cases, what should philosophers do in the case of genuine disagreement?

Kornblith considers Thomas Kelly's answer that we must simply turn to the arguments and believe what they demonstrate, rejecting it for reasons similar to the ones considered above.[4] Pragmatically speaking, a disagreement might well be clarified or even resolved if both parties lay out all the relevant arguments, explicitly asses their cogency and their reasons for finding the arguments cogent, and so on. But since beliefs may still (and often do still) differ after such careful analysis by both parties, the question remains as to how the disputants ought to react to the persistent disagreement.

The way Kornblith develops his own answer to the question is interesting, and worth retracing. His most significant move is to expand the discourse beyond the two-party scenario often presupposed by epistemologists of disagreement, shifting the focus of the discussion to the tenability of dissent within a community. His guiding example here is Kelly's discussion of the Newcomb Problem, introduced into decision theory by Robert Nozick in 1969 and debated ever since.[5] Kelly tells us that when the problem was first introduced, opinion among decision theorists was split more or less evenly, but that over the next three decades consensus shifted in favor of the two-box option. According to Kelly, these facts about the distribution of opinion should have no bearing on which decision we make now; we must stick to the arguments and make the most rational decision possible. For we can always imagine possible worlds in which the distribution of opinion had developed in the opposite way, or where it remained evenly split. In general, the line between possible disagreement and actual agreement is "an extremely contingent and fragile matter."[6] Thus, the lack of disagreement in a field – in a word, consensus – is no guide for making difficult decisions.

Kornblith challenges Kelly on his own grounds. According to Kornblith, if we look at the actual history of formal areas of philosophy like logic and decision theory (along with mathematics), we have every reason to think that consensus in such fields is a reliable guide for making decisions. In their infancy, to be sure,

[4] See Thomas Kelly, "The Epistemic Significance of Disagreement," *Oxford Studies in Epistemology* 1 (2005): 167-196.

[5] Kelly, "The Epistemic Significance of Disagreement," 182-185.

[6] Kelly, "The Epistemic Significance of Disagreement," 181.

these fields gave rise a wide variety of opinions among practitioners. Kornblith notes that the uncertainty correspondent to this diversity extended to arguments as much as it did to results; it wasn't until these fields developed to a certain extent that we were able to recognize what concepts had resonance and what methods worked. But once they did develop, the passage of time brought with it increased sophistication and, concomitantly, increased consensus. From this, Kornblith draws a rather strong conclusion:

> Even among experts, of course, convergence of opinion is no guarantee of truth, but one would have to be a radical skeptic about mathematics, logic, probability and decision theory to think that convergence of opinion is not, at this point in the history of these fields, evidence of truth. And at this point in the history of these fields, I think it is fair to say, radical skepticism is no longer a rational option.[7]

In other words, it is not only irrational to think that mathematics and decision theory do not provide us with truth, it is irrational to think that the bare fact of consensus in these fields does not serve as evidence for their truth. It is not far from this conclusion to the strong normative one that wherever there is a consensus in controversial matters, one ought to believe what the experts believe. Kornblith admits that "convergence of opinion is no guarantee of truth" and encourages us to attend to the relevant arguments or evidence as closely as possible, but despite these apparent concessions, his normative prescription remains the same: whether you are a novice or expert, if there is a consensus, you must go with the consensus.

Contra Kelly, it is of no import that we can imagine other possible worlds where the consensus might have turned out differently. Kornblith argues that we can certainly picture a world where Gödel's incompleteness theorem was not discovered by Gödel or wasn't discovered at all, but it would be difficult to picture a world where Gödel published his theorem and it was rejected by the mathematical community at large. Possible, but difficult, for it would require us to rescind our basic faith in the competence of mathematicians in a way that belies the actual progress made within the mathematical community over time. The simple fact that we can imagine worlds where the experts get everything wrong doesn't have any bearing on the fact that in this world, they tend to get things right over time.

According to Kornblith, the basic normative principle operative in the restaurant bill case is the same in the cases of mathematics, logic, and decision theory. If I go out to dinner with 17 mutually distrustful friends who all want to

[7] Kornblith, "Belief in the Face of Controversy," 40-41.

add up the bill separately, and everyone but me calculates the same sum, the obvious thing for me to do in this case is to accept that I have made a mistake. As Kornblith tells us,

> Things are no different if we move from dividing the bill at a restaurant to solving a problem in decision theory.[8]

Increased complexity doesn't change the fact that the experts know best.

But we ask again: what about philosophy? It's hard to think that Kornblith hasn't strayed from his focus a bit. In fact, he's effectively made his argument already; the case against the epistemic authority of philosophy is made entirely by analogy to fields like mathematics, decision theory, and the natural sciences. If the philosophical community was like the mathematical and scientific communities in the sense that there has been undeniable progress within the community over time, then we would be able to treat philosophers as reliable experts in their field and impart the same epistemic authority to them we do to individual scientists. In that case, where there was a roughly even divide on some controversial philosophical issue, we would be rationally compelled to suspend judgment on the issue. "But," Kornblith tells us,

> surely it is not reasonable to believe that the philosophical community is like the mathematical or scientific communities in relevant respects. We don't have a long history of steady progress on issues, and, as a result, the case for deferring to community opinion is thereby weakened.[9]

And the case for the epistemic authority of philosophers – a case, we note, that Kornblith never bothers to make – is also thereby weakened. For where we see no track record of success in the field, we have no reason to trust the authority of the practitioners within that field.

The conclusions that Kornblith draws from this argument are merciless. Because of the lack of progress in philosophy,

> the only conclusion we can reasonably reach is that there is no basis for opinion here on anyone's part at all.[10]

By Kornblith's standards, it would seem to be as irrational to hold a genuine belief about a philosophical matter – *any* philosophical matter – as it is to question whether consensus in logic is really evidence that logicians are converging on the truth. To hold such opinions would be as presumptuous as it would be to believe that my calculations were correct and those of my 17 dinner companions were

[8] Kornblith, "Belief in the Face of Controversy," 43.
[9] Kornblith, "Belief in the Face of Controversy," 44.
[10] Kornblith, "Belief in the Face of Controversy," 45.

wrong. By all means, Kornblith tells us, we should continue to swap arguments like low-value baseball cards and hold our little conferences to make noise out of our precious beliefs, but only so long as at the end of the day, we remind ourselves that it's all much ado about nothing. Kornblith calls this curious defeatism "epistemic modesty." Others might call it a good reason to not do philosophy.

3. Lessons From the History of Philosophy

A natural first reaction to Kornblith's argument would be to reject the premise that philosophy hasn't made the same progress as the other disciplines discussed. No doubt progress of some kind *has* occurred in math and science, but surely it has transpired in philosophy as well. Suppose we take Plato and Aristotle to mark the beginning of systematic philosophy. Very few philosophers still believe in immaterial Forms that subsist independently of the material world, and even fewer believe that because of the superior constitution of their soul, philosophers ought to rule as kings. Aristotle's defense of the institution of slavery is no longer tenable, nor are his conclusions about the inferiority of women. Aside from some general platitudes about virtue and intellectual rigor, in fact, there is little in the philosophy of the ancient Greeks that many rational people endorse today. Is this not progress?

As appealing as this approach seems on first glance, I think it is perhaps the weakest response to Kornblith. The kind of progress Kornblith has in mind isn't just a matter of there being a different set of beliefs now than there was two millennia ago, or even a century ago. As Kornblith says, even fashion exhibits such shifts in consensus. Progress implies *cumulative* change that converges on a fixed point. But there's certainly no track record of fixed conclusions that we can point to as evidence of progress in philosophy. Even in ethics, the domain that philosophers most often try to reserve for themselves against the encroachment of modern science, what we see isn't linear progress so much as perpetual shifts in disagreement. Philosophers might not fancy themselves to be kings anymore, but they're not much help in telling us who ought to rule instead.

If we accept Kornblith's definition of progress, then, philosophy has exhibited little to none. But where does this definition come from? We know Kornblith's answer: it is simply the name for the successive advance towards truth evident in formal philosophy, math, and natural science. But does it really make sense to talk about progress in logic and progress in physics as if they were the same phenomenon? Philosophers of science like Kuhn and Lakatos have served the valuable function of reminding us that progress in physics is not nearly as clear-cut as we like to think, and there is an important sense in which theories

and observational data are incommensurable from one research program to the next. But suppose we accept for the sake of argument that progress in physics is clear-cut: what relation does the ability to predict and experimentally verify the behavior of natural objects have to progress in logic? Logicians make no predictions and conduct no experiments. There is nothing in the realm of nature that confirms the law of noncontradiction, or demonstrates that the fallacy of ambiguity is a fallacy. The only thing that logical progress and scientific progress seem to have in common is that they are marked by increasing consensus. But if the bare fact of expert consensus can't be proof of progress just on its own, it certainly can't serve to indicate truth either.

Perhaps consensus in logic – and mathematics, for that matter – is nothing more than the inevitable outcome of a large group of intelligent people finding the most agreeable ways to manipulate symbols of their own making. Maybe such a view is mistaken, but no one is going to demonstrate that with an experiment in the lab. If I claim that water flows upwards, on the other hand, a simple experiment could prove me wrong. Physics and logic don't share methods, objects, arguments, or procedures of confirmation and falsification; why should they share the same standard of progress?

The point here isn't just that Kornblith is glossing over some very important details – although he is – but rather that he is strewing loaded terms across such a diverse terrain that they cease to have any useful meaning at all. I doubt that the truth of physics has much in common with the truth of logic, if indeed the latter really gives us what ought to be called truth. Moreover – and this is the important point – even if the sense of truth in those fields was in some important sense congruent, I see no reason at all to think that the further extension to fields like metaphysics or epistemology is valid. Kornblith is committing his own fallacy of ambiguity here, and doing so with one of the most notoriously slippery terms of our language. Philosophers can't even agree on what truth *is*; what gives Kornblith the epistemic authority to throw around the term as if we all, philosophers and scientists alike, had come to a consensus about what it meant?

This lacuna in Kornblith's argument wouldn't be so important if he weren't basing his entire argument on what amounts to an analogy between philosophy and other disciplines. But, as we have emphasized above, this is precisely what he is doing. This move seems especially egregious since, in the same article, Kornblith has stated that disagreement in perceptual or mathematical cases seems importantly different than disagreement in philosophy. In math or science (where conclusions are ultimately supported by reports of perceptual observation), there is typically an accepted method of confirmation that we can turn to in the case of

disagreement. But in philosophy the path from disagreement to resolution is rarely so apparent. I'd happily go a step further: when it comes to philosophy, I'm not so sure that resolutions or solutions always exist.

Here Kornblith might remount his attack. If philosophical problems can't be solved - or, what amounts to the same, if we don't know what it would mean to solve them - then surely philosophy never had any epistemic authority to begin with. What good is a problem that can't be solved, or a theory that can't be confirmed? I think the correct answer is that they do a great deal of good, or in any case, have a significant impact. As any philosopher in our dying western Lyceum will readily tell you, the study of philosophy has all kinds of practical benefits at the introductory level: it sharpens critical thinking skills, encourages close reading, makes us question assumptions and assumptions behind assumptions, and perhaps even casts doubt on the idea that holding a fixed set of beliefs is really the most rewarding and productive way to employ one's rational capacities.

We need not stop at philosophy 101, however. Philosophy also has influence on the world-historical scale. Hegel is typically lambasted as the caricature of an abstruse, hopelessly abstract metaphysician, but without Hegel there would have been no Marx and the twentieth century would have turned out very differently. If we look at Eastern philosophers like Confucius, the 'real world' applications are even deeper and more pervasive. Moreover, philosophy has served as the cradle for virtually every other major intellectual discipline. Without Aristotle, there would have been no political theory, aesthetic criticism, psychology, physics, biology, optics, etc. Without Plotinus, there would be no God as Christians understand that concept in the West today. If Kornblith wants to deny the epistemic authority of philosophers on the ground that they don't do what the scientists do, he would do well to remember that without philosophy there would be no science as we know it.

In sum, Kornblith introduces the dimension of history into the epistemology of disagreement without bothering to make the slightest effort to understand what makes the history of philosophy unique, or valuable, or epistemically significant. If he had done so, he would have realized that philosophy doesn't make progress like physics does because it is a discipline concerned with foundations, and each new epoch in human history brings with it different requirements for such foundations. Cartesian mechanism was both product and mirror of its age, as was Aquinas' scholasticism, Nietzsche's aesthetic anti-nihilism, Quine's holism, and Kornblith's own blithe scientism. One need not be a relativist to understand that different eras in history give rise to different

problems and accept different kinds of answers, and one would be myopic to deny that philosophy has accomplished the task of formulating and propagating such questions for longer and at a deeper level than any other intellectual enterprise launched by humankind. The movement of philosophy does not proceed from point A to point B, but from point A to the ground beneath it; if it so happens that it digs all the way through and ends up on the other side of the world speaking a foreign language, ought we to deny its epistemic authority on the grounds that it didn't take the path we thought it would?

4. Conclusion: Mentats and the Pre-born

The world of *Dune* is a veritable epistemological funhouse. In addition to the Mentats, Herbert gives us a very different – and much stranger – model of superhuman intellectuality. The 'pre-born,' as Herbert names them, are human beings who are exposed to a certain type of drug (the melange spice) while still in the womb and thereupon come immediately to mature consciousness. But that's not all: once exposed to the drug, they are imbued with the complete memories of every ancestor in their genetic line. They carry these memories with them for their entire lives, and as such, are able to draw upon an unimaginably vast reservoir of historical experience to guide them along their own paths. Because they have (vicariously) seen it all before, the pre-born are impeccable judges of human behavior, and exhibit a kind of stoic remove from the complications of everyday life. They don't have the problem-solving prowess of the Mentat, but they have something better: the ability to understand why and how problems arise in the first place, and to judge their significance for the human condition as it stands in their historical moment. Perhaps it is for this reason that it is the pre-born who always end up on the throne, and Mentats who end up serving them.

Few educated people in our own world, aside from radical skeptics and the credulously religious, are likely to deny the epistemic authority of scientists and mathematicians. The ability to function as a Mentat might be beyond the power of any living human, but the factor of time and the contributions of the community make up for this lack: scientific disciplines accomplish collectively and diachronically what the Mentat accomplishes individually and synchronically. I proposed at the start of this essay that the Mentat ideal is also a suitable model for the philosopher. For scientifically-minded, results-oriented thinkers like Kornblith, it might well be. But if my objections against Kornblith in this essay have been successful, then I hope to have laid the ground for a different kind of ideal. Those adhering to this ideal would shift their attention to the historical foundations and precedents for what comes to be. They would pay more attention

to the genetic causes and conceptual roots of contemporary problems than they would to finding the quickest solution to those problems. They would find an altogether more suitable ideal in the pre-born.

This divide between the Mentats and the pre-born might be taken to map onto the divide between analytic and continental philosophy in interesting ways, but within the confines of this essay I can't pursue this idea. My concern is to suggest that although they may adopt different methods, deal with different problems in different ways, and ultimately come up with different results, neither of these ideals intrinsically holds any more epistemic authority than the other.

There's nothing wrong with holding *some* philosophy up to the standards set by modern mathematics and science, so long as the philosophy in question is similar enough to those fields in relevant ways. But there is also nothing wrong with practicing philosophy as an autonomous discipline, judging it by its own standards and in accordance with its own historical successes and failures. If we do the latter, questions of progress and consensus simply don't matter as much as Kornblith thinks they do. The practical virtues of philosophy persist unperturbed in either case, so as far as I can tell there is no normative basis for choosing one over the other. But whichever way we go, thousands of years of endlessly fascinating ideas clamor to be acknowledged, heeded and heard anew. The question is whether we want to hear them.

DEFENDING THE UNIQUENESS THESIS: A REPLY TO LUIS ROSA

Muralidharan ANANTHARAMAN

ABSTRACT: The Uniqueness Thesis (U), according to Richard Feldman and Roger White, says that for a given set of evidence E and a proposition P, only one doxastic attitude about P is rational given E. Luis Rosa has recently provided two counterexamples against U which are supposed to show that even if there is a sense in which choosing between two doxastic attitudes is arbitrary, both options are equally and maximally rational. Both counterexamples work by exploiting the idea that 'ought implies can' and trying to spell out situations in which some inferences are beyond the capabilities of some reasoners. I argue that on a descriptive account of doxastic rationality, questions of whether 'epistemic ought implies can' can be bracketed and that at least one of the inferential moves that Rosa describes in his cases is irrational. I further argue that a descriptive account of doxastic rationality is the appropriate notion of rationality that is to be considered when evaluating U. If my argument for a descriptive account of rationality is successful, then we have reason to revise our use of the term rationality to fit this descriptive understanding.

KEYWORDS: epistemology, uniqueness thesis, epistemic deontology

Introduction

We have many beliefs, and we would like to think, for any given belief, that not only is it true, but also that we are justified in believing this to be the case. However, when we meet others who are just as rational as ourselves and who possess the same evidence, but take a different doxastic attitude, we may be concerned about what we should do, whether one or more of us should or should not revise our beliefs so as to be aligned with that of others'. On the one hand, it might seem that if rational disagreement were possible, then the fact that we have different beliefs given the same evidence, may not be too disquieting. Conversely, it would seem inappropriate to so easily dismiss the concern that perhaps one of the parties has indeed made a mistake while forming his beliefs. After all, the question may arise as to why we hold one particular belief rather than another. To this, the only answer which would be consistent with taking that belief to be fully determined by the reasons we have is that this belief is more appropriate than any other. This could only be the case if something like the Uniqueness Thesis were true.

According to the Uniqueness Thesis:

> Uniqueness Thesis (U): 'Given a body of evidence, [at most one doxastic attitude[1]] is the rationally justified one'[2]

Roger White defends U by showing that alternatives to U are untenable. White calls all these alternatives 'permissive theses' and defines them as follows. The first one is Extreme Permissiveness (EP):

> (EP): 'There are possible cases in which you rationally believe P, yet it is consistent with your being fully rational and possessing your current evidence that you believe not-P instead'[3]

Luis Rosa defends EP by offering a counterexample in which either P or not-P is allegedly permissible if the relevant evidence contains, unknown to oneself, inconsistent premises.[4]

The second is Moderate Permissiveness, (MP) which does not sanction such vastly opposed doxastic attitudes, but still sanctions a narrower range of such attitudes. Rosa presents one version:

> (MP): 'There are cases in which a certain body of evidence E makes believing that P rational, but E could also make suspending judgment about P rational'[5]

Rosa defends MP by offering a counterexample to U in which both suspending judgment and believing P are permissible, if in one case, an agent is incapable of making a particular inference.[6] Of course, the leeway in doxastic attitudes sanctioned by MP may be much smaller.[7]

In this paper, I will defend the Uniqueness Thesis against Rosa's counterexamples,[8] by appealing to a non-deontic account of rationality. I shall specifically defeat the counterexamples that defend some version of either EP or

---

[1] To be clear, I am only concerning myself with the degree of confidence that a person can have with respect to a particular proposition. I do not take up whether it is permissible to use a given proposition as a premise in subsequent reasoning nor am I, in this paper, concerned with pragmatic encroachment.

[2] Richard Feldman, "Reasonable Religious Disagreements," in *Philosophers Without Gods,* ed. Louise M. Antony (New York: Oxford University Press, 2007), 205.

[3] Roger White, "Epistemic Permissiveness," *Philosophical Perspectives*, 19, *Epistemology* (2005): 447.

[4] Luis Rosa, "Justification and the Uniqueness Thesis," *Logos & Episteme* III, 4 (2012): 572-574.

[5] Rosa, "Justification and the Uniqueness Thesis," 572.

[6] Rosa, "Justification and the Uniqueness Thesis," 574-575.

[7] White, "Epistemic Permissiveness," 453.

[8] Rosa, "Justification and the Uniqueness Thesis," 571-577.

MP. If no counterexample to U can be constructed, it may be because U is necessarily true.

Motivating U

By definition, if a given belief B is based on extremely permissive evidence, E, then one could have rationally based the opposite belief on E. White's objection to EP is that if this is the case, then inferring B from E is no better than using an arbitrary process like taking a belief inducing pill or flipping a coin. However, if one's belief is arrived at via such an arbitrary process, it is irrational. The point about arbitrariness is crucial as a body of evidence, as a whole, cannot be said to support a proposition if it does not support the proposition more than its negation. It could be said that MP exhibits this same arbitrariness, but only to a lesser degree; only that the intuitive force of the point is not as strong as with EP. According to Kelly,[9] in cases where the permissible range of views is so narrow that all permissible doxastic attitudes are only minutely different from one another, choosing one of those attitudes over another does not seem viciously arbitrary.

One way of resisting U is to show via counterexamples, that having multiple doxastic attitudes from E is not arbitrary. Another strategy is to show via counterexamples, that those multiple doxastic attitudes, even if arbitrary, are not irrational. That is to say, that even if reasoning from the same piece of evidence leads different people to different conclusions, they may still be instantiating good types of reasoning. The focus of this paper will be the second strategy.

This second approach[10] is to invoke certain deontic features of reasoning, namely the idea that 'ought implies can,' that are prominent in internalist accounts of epistemic justification.[11] Implicit in this strategy is the idea to say that some inference is rational implies that such an inference is either obligatory or permissible. Since real reasoners are incapable of thinking about all of their beliefs at the same time, they may fail to notice inconsistent beliefs that they may hold even when they make a sincere attempt to verify that all their beliefs are consistent with one another. I will call this way of using the term 'rational' d-

[9] Thomas Kelly, "Can Evidence Be Permissive?" in *Contemporary Debates in Epistemology*, eds. Ernest Sosa, Matthias Steup, and John Turri (Oxford: Wiley-Blackwell, 2013): 298-311.

[10] See Anthony Booth and Rik Peels, "Why Responsible Belief is Permissible Belief," *Analytic Philosophy* 55, 1 (2014): 75-88 and Rosa, "Justification and the Uniqueness Thesis," 571-577.

[11] Richard Foley, "Conceptual Diversity in Epistemology," *The Oxford Handbook of Epistemology*, ed. Paul K Moser (Oxford: Oxford University Press, 2002), 179-180

rational[12] and contrast it with another way of using the term rational. If a given inference is beyond the capabilities of a reasoner, the inference is neither obligatory nor even permissible for that reasoner. If that is the case, then a person may have done the best that she was capable of, given her abilities, and still make different inferences from the same body of evidence due to chance or some other factors. For instance, limited reasoners may have to make tradeoffs between acquiring true beliefs and avoiding false beliefs. Different attitudes towards epistemic risk would make different inferences rational for different people. Since such a person could not have done more, she is not obligated to do more and cannot be called d-irrational. Since she has not failed her epistemic duties and even if she did have an epistemic duty, her failure to do so could not have been avoided; she is beyond criticism on this account. As such, both doxastic attitudes towards the evidence are rationally justified, though, perhaps, not simultaneously. Often, in such counterexamples, details about which inferences are made are not available. In such cases it becomes impossible to affirm or deny whether any of the reasoners' inferences are rationally defective. Luis Rosa[13] presents counterexamples which are more detailed in this respect. When the exact inferences made are supplied, it is possible to assess the counterexample to see if Rosa is indeed correct about the rationality of both reasoners.

If d-rationality was indeed the relevant way in which proponents of U think there can only be one rational response to the total evidence, then U is in trouble. I will argue, therefore, that there cannot be more than one rational response. I contend that there is another, non-deontic way in which we use the term 'rational' which is appropriate in at least some situations and more importantly, is the relevant sense in which U is meant and in which U is true.

One way of talking about rationality which, arguably, cannot have deontic connotations is when we talk about defects in a person's reasoning.[14] One of the counterintuitive implications about using d-rationality in such a situation is that

[12] D-rationality as such seems to assume some significant level of doxastic voluntarism. I am, for the purposes of this paper, willing to grant this assumption. If this assumption seems unfounded, then that is one more reason to think that d-rationality is the wrong way to talk about epistemic rationality. Objections to U which rely on d-rationality being the right account of epistemic rationality will likewise be unsuccessful.

[13] Rosa, "Justification and the Uniqueness Thesis," 571-577.

[14] See Gideon Rosen, "Nominalism, Naturalism, Epistemic Relativism," *Philosophical Perspectives*, 15, *Metaphysics* (2001): 85, Booth and Peels, "Why Responsible Belief is Permissible Belief," 79, and David M. Holley, "Religious Disagreements and Epistemic Rationality," *International Journal of the Philosophy of Religion* 74 (2013): 41, for examples of such usage.

to say that some instance of reasoning is without rational defect is to ordinarily imply that it is perfect. Consider the example of a person, Amanda, who is unable to use *modus tollens*.[15] While we may in some circumstances be unwilling to call her irrational when she fails to use *modus tollens* to arrive at what seems to us a fairly obvious conclusion, we would be hesitant to say that her reasoning on the issue lacked rational defect. Since, as according to Rosa's description of the case, the inference that Amanda failed to make would have been obvious to us if we were in those circumstances even if it was not so to her, it seems intuitively obvious that her failure in this case is rationally defective. Yet, per definition, since she was unable to use *modus tollens*, she had no obligation to do so and her reasoning thus lacked any d-rational defect.

In order to account for the above intuition, the term 'rationality' must be used in some other sense in which either 'ought' does not imply 'can' or in which it is not the case that one ought to make an inference which is rational. In order to bracket the question of whether 'ought implies can' applies to the epistemic ought, I will opt to use the term 'rationality' in the second way, namely, by taking rationality to not imply 'ought' or even 'may.' Then, even if 'ought' does imply 'can,' it does not apply when I use the term rational in this second sense. For clarity, I will use the term 'p-rational' to refer to the sense in which 'rational' is used in these cases. The question of whether a given inference is obligatory or permissible can be evaluated separately.

We may say that an inference is p-rational if and only if no performance errors are made. A person makes a performance error in reasoning about a proposition P if the inference she makes is invalid or if she fails to make an inference that would have been valid had it been made and that inference bears on whether P. After all, if making that valid inference would not have resulted in a different doxastic attitude about P there are at least some cases where it would be unclear if such an error was relevant or significant.[16] The preceding account of performance errors is not necessarily complete. Moreover, it is so demanding that all of us can be expected to fall short of its demands most of the time. However, this should not surprise us. Except in particularly simple reasoning tasks we cannot expect to avoid entirely any performance errors in our reasoning. Moreover, since 'ought implies can' does not apply to p-rationality, the fact that it

[15] Rosa, "Justification and the Uniqueness Thesis," 574-575.

[16] There are also some cases where such a failure would be relevant but I need not explore those instances in order to make my point in this paper. See John Turri, "On the Relationship between Propositional and Doxastic Justification," *Philosophy and Phenomenological Research* LXXX, 2 (2010): 312-326, for an excellent discussion of some of these cases.

is a demanding account of rationality is not a criticism that is applicable. After all, demanding-ness matters only if 'ought' implies can. It is adequate for my purposes if making an invalid inference as well as the failure to make a valid inference are performance errors. However, how do we characterise a valid inference? For the purposes of this paper, it is adequate to note that *modus ponens* and *modus tollens* are valid inferences as is replacing a conditional (If P then Q) with the disjunction (not-p or Q) containing the negated antecedent.

I will now use this account of p-rationality to demonstrate that the alleged counterexamples to U, provided by Rosa that attempt to show that there is more than one rational doxastic response to some given evidence, are unsuccessful. If it can be shown that one or more of those inferences are invalid, then it would disqualify those inferences from being rational.

Rosa's defence of Extreme Permissiveness (EP)

To defend EP, Rosa uses the example of Michelle who draws conclusions from a body of evidence E which, unknown to her, contain inconsistent premises. Michelle believes about George, the following:

1. George is tired, but willing.
2. If George is tired he will rest or sleep.
3. George is not willing or it is not the case that he will rest or sleep.[17]

From the first and second premises, Michelle can conclude that George will rest and (or) sleep. From the first and third, she can conclude that he will not. Given Rosa's description, it should be obvious to the reader that the three premises are inconsistent even if it is much less apparent to Michelle. Abstracting from the particular content of the beliefs in question, the case can be described in the following way:

> Michelle could, using some of her premises found in E, rationally infer a conclusion C. Michelle could also rationally infer from some other premises found in E that not-C.

Since, according to Rosa, either of the inferences would be rationally warranted, and Michelle, by stipulation and given the circumstances, could not have known that E contained inconsistent premises, both C and not-C are rational inferences from E for Michelle. Rosa further bolsters this conclusion by appealing to the intuition that we would be reluctant to call Frege irrational just because he

[17] Rosa, "Justification and the Uniqueness Thesis," 572.

failed to notice the contradictions inherent in his logical axioms. By such an exacting standard, Rosa argues, none of us are rational.[18]

This bullet is a tough one to bite, only if the appropriate sense of rational used is d-rational. But, if p-rationality is the appropriate notion of rationality in this situation, it is not so difficult to think that reasoning from inconsistent premises is not rational even if it was unknowingly done. However, how could this be the case? The permissivist may argue that even if there is still a mistake made in reasoning from inconsistent premises, it is not a performance error as no invalid inferences are made, and only a performance error can implicate the quality of a given inference. Consider a variant of the above situation where Michelle has misleading evidence E' instead. E' is consistent, but at least one of its premises, unknown to her, is wrong. There is further nothing that Michelle could reasonably have done to figure out that E' contained a false premise. In this case, there is no sense in which Michelle is irrational when she derives C' by applying *modus ponens* to E'. Likewise, there is no performance error when carrying out *modus ponens* even when the premises are false. Michelle might be mistaken in thinking that E' was appropriate to base her reasoning on, but we cannot necessarily attribute any performance error on her part. After all, if E' had been true, using *modus ponens* on E' would necessarily have yielded a true conclusion.

If one's premises are inconsistent, there are valid inferences that one could make from one's premises that would show that this is indeed the case. The only way in which one could be unaware that this is the case is if one failed to make a particular valid inference from one's own premises. Even if this lapse is excusable, unavoidable and thus often committed by all human reasoners, it is still a performance error. Thus, a person makes a performance error in failing to notice that she has inconsistent premises.

A permissivist may still argue that even if the failure to infer that E contains inconsistent premises is some sort of performance error, it is not necessarily a performance error that is relevant to whether C or not-C. Let me illustrate: If there is only one pair of inconsistent premises one of which is true, E can be made consistent by eliminating the one premise, A, to yield E' or eliminating the other one, not-A, to yield E''. Presumably, C can be validly inferred from E' and not-C from E''. Even if the fact that E contains inconsistent premises was inferred, since either A or not-A could be rejected to make the set of beliefs consistent, it seems that not only is either way of resolving the contradiction equally rational, identifying that a contradiction exists still leaves us in a position of having to decide between inferring C and not-C via the intermediate step of resolving the

[18] Rosa, "Justification and the Uniqueness Thesis," 573-574.

contradiction one way or the other. If identifying a contradiction necessarily resulted in one conclusion being identified as more rational than the alternative, the failure to identify the contradiction would be a relevant performance error. But, the permissivist could argue that this is not the case here since there is, according to the permissivist, no uniquely best way of resolving the contradiction.

However, the permissivist is mistaken about this last claim. Identifying the existence of a contradiction is relevant because there is a uniquely rational way of resolving contradictions. In addition to rejecting A or not-A, there is also the option of suspending judgment on both premises to yield E'''. It is the case that either A has more support, less support or as much support as not-A. Let us suppose that Michelle possesses evidence T bearing on A or not-A. If the total evidence T supports either A over not-A or vice versa, then one or the other is to be rejected yielding either E' or E''. Given that one of the propositions enjoys greater support than the other, Michelle would be making a performance error if she rejected the proposition which enjoys greater support. Thus, if one of the propositions had more support than the other, there would be only one uniquely best response to the evidence E.

However, what if both A and not-A were equally well supported by T? Might it not be the case that rejecting either one of A or not-A was equally good? If rejecting A was just as good as rejecting not-A, suspending judgment about A or not-A would be better than rejecting either proposition. This is easily demonstrable once we translate the degree of support for the statement A provided by T as the probability that A given T, p(A|T).

Since, per assumption, T supports A and not-A to the same extent,

$$p(A|T) = p((\text{not-}A)|T) \text{ ---------- (1)}$$

Since that A and not-A cannot both be true at the same time,[19] the probability axioms cannot be violated. Since A and not-A are mutually exclusive and exhaust all logical possibilities,

$$p(A|T) + p((\text{not-}A)|T) = 1 \text{ ----------(2)}$$

Solving for (1) and (2),

$$p(A|T) = p((\text{not-}A)|T) = 0.5$$

Since, according to the evidentialist, epistemic rationality just amounts to apportioning one's degree of belief in a proposition to the support for that proposition by the evidence that one has, suspending judgment about A and not-A is the only response that comes closest to being proportional to the support for

[19] For the purposes of this paper, I will exclude non-standard logics.

those propositions by the total evidence T. Thus, any response other than suspending judgment about A and not-not-A would be a performance error.

I have thus shown that knowingly reasoning from inconsistent premises constitutes a performance error. EP is thus, absent a better counterexample, indefensible, at least when we use p-rationality. Having disarmed Rosa's EP counterexample with p-rationality, I will now turn to Rosa's counterexample defending MP.

Rosa's Defence of Moderate Permissiveness (MP)

The argument against MP is equally strong. Rosa defends MP by appealing to the example of Amanda who in possible world W1, is able to infer that she did not press a button from the following premises via *modus tollens*:

1. If Amanda presses the button, her computer will get infected with a virus.
2. Amanda's computer is not infected with a virus.

Amanda has a counterpart in possible world W2 who is incapable of carrying out *modus tollens* and thus suspends judgment about whether she pressed the button. Moreover, since she does not believe she has good reason to believe that she pressed the button, according to Rosa, it would be absurd to say that she is irrational in suspending judgment about whether she pressed the button. Thus, if Rosa is right, both believing and suspending judgment about whether she pressed a button are fully rational responses to E in at least this case.[20]

However, the same criticism that was levelled against the EP case can also be levelled against the MP case. It may be that Amanda$_2$ is not required to use *modus tollens* if she cannot and if she sees no reason to. Or even if she is required to, perhaps she is blameless because she lacks the capacity to do so. Yet, it is obvious that Amanda$_2$'s inability to use *modus tollens* means that she is pre-disposed to fail to make many valid inferences from her premises. This would constitute a performance error on Amanda$_2$'s part and she would not be p-rational. In fact, it is hard to say that Amanda$_2$ possesses even a normal reasoning ability let alone a particularly good one. The only way out for Rosa here is if it were somehow the case that *modus tollens* was not valid in W2 but valid in W1. But this seems implausible. In both W1 and W2, using *modus tollens* on true premises necessarily gives us true conclusions. This means that modus tollens has to be valid in both worlds.

[20] Rosa, "Justification and the Uniqueness Thesis," 574-575.

It thus seems to be the case that Rosa's defence of MP is unsuccessful. If no better counterexample is available, MP is indefensible. Nevertheless, there might be other counter-examples to U that are even more refined than Rosa's. These counterexamples would have to unambiguously show instances where more than one conclusion can be inferred from a given body of evidence and that both inferences are impeccable.

Conclusion

The above discussion has rebutted counterexamples aiming to establish that even if two doxastic responses to some evidence are arbitrary, they can still be rational. In order to establish U, I need to rebut counterexamples that purport to show that sometimes, multiple responses to some given evidence are non-arbitrary. Even if I end up showing that U is true given p-rationality, the permissivist might still claim that using the word rationality to mean p-rationality is only appropriate in some circumstances and d-rationality fits better with the way we speak. Even if as a descriptive matter, this was true; there might be some reasons to revise the way we use the word rational to refer to p-rational instead. For one, anything we can say in terms of d-rationality can be said in terms of p-rationality. Since d-rationality refers to what inferences we are obligated or permitted to make, the set of d-rational inferences can be cashed out in terms of the set of inferences that most closely resemble p-rational inferences and which the reasoners in question are capable of making.

On the other hand, there are some things that can be expressed in terms of p-rationality which cannot be expressed in terms of d-rationality. Suppose there are a set of inferences all of which are beyond the capabilities of a given agent, some of which contain fewer performance errors than others. For instance, suppose a large set, S, of statements exists such that some of them contradict one another. S also contains an indicative conditional C for which the consequent is also contained in S. Because S is very large, John, an ordinary but competent reasoner would not be able to identify that S contains inconsistent premises. It is also the case that because S is very large, John would not be able to affirm the consequent with respect to C. Since both inferences are beyond John's capabilities, it is not d-rational to make either inference. Thus, if d-rationality is the concept that is referred to by doxastic rationality, both inferences are on par. However, if p-rationality is the relevant notion of rationality, identifying a contradiction is clearly superior to affirming the consequent.

Thus p-rationality cannot be cashed out in terms of d-rationality. This asymmetry means that using rationality to refer to p-rationality is more conducive

to analytic clarity. This also means that it is in terms of p-rationality that the Uniqueness Thesis should be evaluated. Evaluating decision procedures for instances of peer-disagreement would be derivative of this.

A second reason to think that p-rationality is the right way to use rationality in this instance is that p-rationality connects up more directly with the motivating intuition behind the Uniqueness Thesis. If all I am concerned about is whether I have violated my epistemic duties in some blameworthy manner, then meeting an epistemic peer who shares my evidence but who still disagrees with me is not necessarily troubling. Having done all I reasonably could have been expected to do in order to ensure that my beliefs are true, I would have no further epistemic duties. However, if I am concerned with acquiring true beliefs and avoiding false beliefs, then the presence of someone who is my epistemic peer and who shares my evidence, but disagrees with me is more worrying. More precisely, the worry is that since both beliefs cannot be true at the same time and since the evidence is, by assumption, true, at least one person has made a performance error in reasoning and that person may very well be me! If this is the case, then any attempt to justify epistemic permissiveness by appealing to our cognitive limitations is mistaken.[21]

[21] I would like to thank Axel Gelfert, Tang Weng Hong, Chetan Cetty, Yeo Shang Long, Peter Kung and Luis Rosa for instructive comments on the paper.

DIFFERENCE-MAKING AND EASY KNOWLEDGE: REPLY TO COMESAÑA AND SARTORIO

Erik J. WIELENBERG

ABSTRACT: Juan Comesaña and Carolina Sartorio have recently proposed a diagnosis of what goes wrong in apparently illegitimate cases of 'bootstrapping' one's way to excessively easy knowledge. They argue that in such cases the bootstrapper bases at least one of her beliefs on evidence that does not evidentially support the proposition believed. I explicate the principle that underlies Comesaña and Sartorio's diagnosis of such cases and show that their account of what goes wrong in such cases is mistaken.

KEYWORDS: easy knowledge, bootstrapping, Juan Comesaña, Carolina Sartorio

Comesaña and Sartorio have recently proposed a novel and elegant diagnosis of what goes wrong in apparently illegitimate cases of 'bootstrapping' one's way to excessively easy knowledge.[1] Here I explain their diagnosis of such cases and show that it is mistaken.

Their analysis of easy knowledge cases rests on the following principle:

> E evidentially supports P (given a background of evidence B) only if the absence of E does not evidentially support P (given the same background B).[2]

They advance this principle in an attempt to formulate a plausible and precise version of the intuitive but somewhat vague thought that a given body of evidence justifies a given belief only if the evidence makes the right sort of difference with respect to the belief. Their proposal is that the evidence makes the right sort of difference to the belief only if the absence of the evidence does not also support the belief. Their explication and defense of the principle make it clear that 'the absence of E' in the principle means the negation of E.[3] Thus, the principle can be stated more precisely this way:

[1] Juan Comesaña and Carolina Sartorio, "Difference-Making in Epistemology," *NOUS* 48, 2 (2014): 368-87.

[2] Comesaña and Sartorio, "Difference-Making," 374.

[3] Comesaña and Sartorio, "Difference-Making," 376.

> Epistemic DM2: E evidentially supports P (relative to a background of evidence B) only if it's not the case that not-E evidentially supports P (relative to B).[4]

As Comesaña and Sartorio note, Epistemic DM2 "gives a non-reductive condition on evidential support."[5] The right-hand side of the principle employs the concept of evidential support and hence the principle "cannot be used in the traditional project of defining propositional knowledge in wholly non-epistemic terms."[6] Accordingly, Comesaña and Sartorio seek to motivate the philosophical usefulness of Epistemic DM2 by making the case that it can shed light on cases of easy knowledge. They provide the following example of such a case: Imagine a subject, Roxanne, who looks at an object in front of her "with her well-functioning eyes open, in a well-lit environment, in the absence of factors that would make the object look a different color than it is" and employs the following reasoning:[7]

> Reasoning A
>
> (1) The object in front of me looks red.
>
> So, (2) the object in front of me is red.
>
> Therefore, (3) my color vision didn't deceive me this time.

Some epistemological theories appear to imply that Roxanne can come to know (3) just on the basis of (1) and furthermore that by engaging in similar reasoning on a variety of occasions while looking at various-colored objects, Roxanne can build a solid inductive case for the reliability of her color vision, thereby arriving at knowledge that her color vision is reliable. Many philosophers find such a procedure to be illegitimate and hold that any epistemological theory that licenses such a procedure is flawed.[8]

---

[4] Or: relative to a given background of evidence, it's not the case that: (E evidentially supports p and not-E evidentially supports p). "Epistemic DM2" is Comesaña and Sartorio's label for the principle.

[5] Comesaña and Sartorio, "Difference-Making," 378.

[6] Comesaña and Sartorio, "Difference-Making," 378.

[7] Comesaña and Sartorio, "Difference-Making," 380.

[8] Reliabilism is a popular target of this sort of objection; see, e.g., Jonathan Vogel, "Reliabilism Leveled," *The Journal of Philosophy* 97, 11 (2000): 602-23. Stewart Cohen argues that any epistemological theory that has what he calls a "basic knowledge structure" is susceptible to the problem of easy knowledge. A distinctive feature of such theories is that they reject the principle that a potential knowledge source K can yield knowledge for S only if S knows K is reliable; see Stewart Cohen, "Basic Knowledge and the Problem of Easy Knowledge," *Philosophy and Phenomenological Research* 65, 2 (2002): 309-29.

Comesaña and Sartorio's aim is not to defend any particular epistemological theory against the problem of easy knowledge but rather to "explain *what exactly is wrong* with cases of easy knowledge, or with a theory that allows that we can come to know 3 just on the basis of 1."[9] They argue that Epistemic DM2 implies that (1) does not evidentially support (3) and that is why Roxanne cannot know (3) on the basis of (1).[10] To establish the claim that Epistemic DM2 implies that (1) does not evidentially support (3), Comesaña and Sartorio argue that if (1) evidentially supports (3), then not-(1) also evidentially supports (3) – which is incompatible with Epistemic DM2. To make the case that if (1) evidentially supports (3) then not-(1) does as well, they suggest that the following line of reasoning is just as plausible as Reasoning A:

Reasoning B

Not-(1): it's not the case that the object in front of me looks red.

So, not-(2): it's not the case that the object in front of me is red.

Therefore, (3) my color vision didn't deceive me this time.

If, as they claim, Reasoning B has approximately the same plausibility as Reasoning A, then the assumption that Reasoning A is plausible and hence (1) evidentially supports (3) yields the result that Reasoning B is plausible and hence not-(1) also evidentially supports (3). Since Epistemic DM2 excludes the possibility that both (1) and not-(1) evidentially support (3), Epistemic DM2 implies that (1) does not evidentially support (3). On Comesaña and Sartorio's view, then, what is illegitimate about the sort of bootstrapping that occurs in cases of easy knowledge is that the bootstrapper bases at least one of her beliefs on evidence that does not evidentially support the proposition believed, and the defect in any epistemological theory that licenses such bootstrapping is that the theory runs afoul of Epistemic DM2.

I think that this diagnosis of what goes wrong in cases of illegitimate bootstrapping is mistaken. While the particular case that Comesaña and Sartorio discuss does involve a violation of Epistemic DM2, there are other cases of apparently illegitimate bootstrapping that do not. It turns out that the fact that the color vision case violates Epistemic DM2 is a consequence of specific features of that case that are absent in other paradigmatic cases of easy knowledge. To see this, consider a version of Vogel's Roxanne case.[11] Suppose that Roxanne has no

[9] Comesaña and Sartorio, "Difference-Making," 380.

[10] Comesaña and Sartorio, "Difference-Making, "373-4.

[11] Vogel, "Reliabilism Leveled," 602-23.

idea whether her car's fuel gauge is reliable. She checks the fuel gauge and reasons as follows (where 'X' indicates some precise level of fuel in the tank – e.g completely full, 5/8 full, 1/3 full):

> Reasoning C
>
> (4) The gas gauge indicates X.
>
> So, (5) the gas tank is X.
>
> Therefore, (6) on this occasion, the reading on the gauge corresponded exactly to the amount of gas in the tank.

As before, Roxanne can use the same sort of reasoning on multiple occasions to build a solid inductive case for the reliability of the gas gauge. So, suppose she draws on multiple instances of reasoning C to arrive at:

> (7) The gas gauge is very reliable – it's disposed to indicate the level of fuel in the tank with a high degree of accuracy.

Suppose that (4) evidentially supports (7) in the way just sketched. If Comesaña and Sartorio's diagnosis of what goes wrong in cases of easy knowledge is correct, then it must be the case that if (4) evidentially supports (7) then not-(4) also evidentially supports (7). In the case of Roxanne and her color vision, Comesaña and Sartorio employed the strategy of generating a line of reasoning parallel to Roxanne's reasoning that began with the negation of Roxanne's first premise. Applying the same strategy to the present case of Roxanne and the gas gauge yields the following line of reasoning:

> Reasoning D
>
> Not-(4): it's not the case that the gas gauge indicates X.
>
> So, not-(5): it's not the case that the gas tank is X.
>
> Therefore, (6) on this occasion, the reading on the gauge corresponded exactly to the amount of gas in the tank.

The conjunction of not-(4) and not-(5) does not imply (6). Suppose that Roxanne is working with data like the following: on Monday, the gauge didn't indicate 5/8 full and the tank wasn't 5/8 full. On Tuesday, the gauge didn't indicate 1/2 full and the tank wasn't 1/2 full. On Wednesday, the tank didn't indicate 1/3 full and the tank wasn't 1/3 full. These are cases in which the gauge got things right within a very wide margin of error; no number of cases of that sort support the conclusion that the gauge has a high degree of accuracy across a range of situations. Therefore, the strategy that Comesaña and Sartorio employed in the color vision case does not work in the gas gauge case and the claim that if

(4) evidentially supports (7) then not-(4) also evidentially supports (7) is unsupported. Consequently, it appears that Epistemic DM2 fails to capture what is wrong with Roxanne's bootstrapping in the gas gauge case.

Furthermore, reflection on this case reveals that the case of Roxanne and her color vision can be modified so that it also does not involve a violation of Epistemic DM2.[12] Consider the following modified versions of Reasoning A and Reasoning B from that case, where 'X' indicates some color:

> Reasoning A'
>
> The object in front of me looks X.
>
> So, (2) the object in front of me is X.
>
> Therefore, (3') my color experience of the object exactly matched the actual color of the object in this case.
>
> Reasoning B'
>
> Not-(1): it's not the case that the object in front of me looks X.
>
> So, not-(2): it's not the case that the object in front of me is X.
>
> Therefore, (3') my color experience of the object exactly matched the actual color of the object in this case.

While the conjunction of (1) and (2) implies (3'), the conjunction of not-(1) and not-(2) does not imply (3'). Consequently, the assumption that (1) epistemically supports (3') does not imply that not-(1) also epistemically supports (3') and so Epistemic DM2 fails to capture what is wrong with Reasoning A'. Therefore, Epistemic DM2 fails to pinpoint just what is illegitimate about paradigmatic cases of easy knowledge.[13] Of course, none of this shows that

[12] Thanks to Justin Snedegar for pointing this out.

[13] An anonymous referee suggests that there is a principle similar to Epistemic DM2 that handles the two cases of easy knowledge that I have argued Epistemic DM2 cannot handle. The modified principle is: E evidentially supports P only if it's not the case that every alternative to E supports P. While it may not always be obvious what the alternatives to a given proposition are, in the gas gauge case and the modified color vision case the alternatives are intuitively apparent. In the modified color vision case, if E = the object in front of me looks red, then the alternatives to E are: the object in front of me looks blue, the object in front of me looks green, and so on. However, this modified principle succumbs to counterexamples. For example, in the color vision case let E = the object in front of me looks red and let P = I'm currently having a visual experience. Intuitively, E evidentially supports P, yet the modified principle implies that E does not evidentially support P since each of E's alternatives also supports P – i.e. that the object in front of me looks blue evidentially supports the proposition that I am currently having a visual experience, that the object in front of me looks green evidentially supports the

Epistemic DM2 is false. For all I have said here, Epistemic DM2 may capture an important necessary condition on epistemic support. However, as I noted above, Comesaña and Sartorio offer Epistemic DM2's usefulness in diagnosing what goes wrong in easy knowledge cases as its main philosophical pay-off. But if I am right, then if we want an adequate explanation of what is wrong with cases of easy knowledge, we should look elsewhere.[14]

proposition that I am currently having a visual experience, and so on. (Note that Epistemic DM2 is not susceptible to this counterexample because not-E (the proposition that it's not the case that the object in front of me looks red) does not evidentially support the proposition that I'm currently having a visual experience).

[14] For helpful discussion of the ideas discussed in this paper, I thank Juan Comesaña and Carolina Sartorio. I presented these ideas at a meeting of the Evidence, Justification and Knowledge reading group associated with the Arche Centre at the University of St. Andrews; I thank the other members of that group for their feedback. Jessica Brown and Justin Snedegar read an earlier draft of this paper and made helpful suggestions, for which I am also grateful.

NOTES ON THE CONTRIBUTORS

Brent Ables is a PhD candidate at the University of Guelph in Ontario. His doctoral research focuses on problems of time in contemporary continental philosophy, but he has a general interest in the history of philosophy and the fate of the discipline in a scientific age. He has presented work on Plato, Nietzsche, Deleuze, and the philosophy of music. Contact: ablesb@gmail.com.

Muralidharan Anantharaman is a recent MA graduate from the National University of Singapore. His main areas of interest are in political philosophy, moral philosophy, meta-ethics, epistemology, social epistemology and decision theory. The bulk of his philosophical research has until now been concerned with Rawls's Original Position, the justification of its features, and various epistemological problems with the use of reflective equilibrium in moral and political philosophy. His thesis can be accessed online at the following URL: http://scholarbank.nus.edu.sg/handle/10635/77767. Contact: amurali284@ gmail.com.

Dionysis Christias received his PhD from the Department of History and Philosophy of Science at the University of Athens. His research interests include epistemology, metaphilosophy, the debate between naturalism and normativism, the philosophy of Wilfrid Sellars and post-Sellarsian philosophy. His most recent articles include "Somatic Intentionality Bifurcated: A Sellarsian Response to Sachs' Merleau-Pontyan Account of Intentionality" (forthcoming in *International Journal of Philosophical Studies*, 2015), "Sellars, the Myth of the Categorial Given, and Correlationism: A Sellarsian Critique of Meillassoux's Arguments against Correlationism" (forthcoming in *Journal of Philosophical Research*, 2015), "A Sellarsian Approach to the Normativism-Antinormativism Controversy" (*Philosophy of the Social Sciences*, 2014), and "An Interpretation and Extension of Sellars' Views on the Epistemic Status of Philosophical Propositions" (*Metaphilosophy*, 2014). Contact: dxristias@gmail.com.

Robert Hudson received his Ph.D. in the Philosophy of Science from the University of Western Ontario in 1992. He has taught at a number of universities throughout North America, and since 2001 has been at the University of Saskatchewan. He works mainly in the areas of epistemology and the history and philosophy of science. His book *Seeing Things: The Philosophy of Reliable*

Observation was published in 2013 by Oxford University Press. Contact: r.hudson@usask.ca.

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, Forthcoming). He is currently co-editing, with Ted Poston, *Best Explanations: New Essays on Inference to the Best Explanation* (Oxford University Press, Forthcoming). Contact: mccain@uab.edu.

Mihai Rusu is postdoctoral fellow in the SOP HRD/159/1.5/S/133675 Project, Romanian Academy, Iași Branch. He received his Ph.D in Philosophy from Babeș-Bolyai University in Cluj-Napoca, Romania (2012). His main research interests include modal epistemology and metaphysics, general epistemology, philosophy of language, and modal logic. Contact: miruss_ab@yahoo.com.

Ju Wang is a PhD candidate at the University of Edinburgh, after having finished his Master degree in Fudan University, China. His research interests are radical scepticism and transcendental arguments. His most recent article is "Closure and Underdetermination Again" (*Philosophia*, 2014). Contact: virgilwang0410@ gmail.com.

Erik J. Wielenberg is Professor of Philosophy at DePauw University in Greencastle, Indiana, and for 2014-15 is a Visiting Fellow at the Centre for Ethics, Philosophy and Public Affairs at the University of St. Andrews. He works mainly in ethics and the philosophy of religion and is particularly interested in issues at the intersection of those two fields. His latest book is *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism* (Oxford University Press, 2014). Contact: ewielenberg@depauw.edu.