Towards a Defense of Moderate Naturalism

by

Nathan Cockram

**A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of**

Master of Arts

in

Philosophy

**Carleton University
Ottawa, Ontario**

**© 2013
Nathan Cockram**

Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-94612-1

Our file Notre référence

ISBN: 978-0-494-94612-1

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

In what follows, I shall argue in favour of Moderate Naturalism, which is a metaepistemological position advocating the compatibility of naturalism and a priori methods of conceptual analysis. More specifically, I have two aims: first, to undermine objections to the compatibility of the relevant doctrines advanced by Hilary Kornblith and Laurence Bonjour; and second, to suggest that Moderate Naturalism is, in fact, superior to the metaepistemological alternatives offered up by Kornblith and Bonjour.

To Mom and Dad.

Acknowledgments

I am especially grateful to my supervisor, David Matheson, for his superb guidance and suggestions. Over the past two years, he has been nothing short of a mentor to me. I would also like to thank Jordan Dodd, Gordon Davis and Matthew Scarfone for all of the productive conversations.

Table of Contents:

Chapter 1: Background	7
1.1 Introduction	7
1.2 Three Background Doctrines	8
1.2.1 The A Priori	8
1.2.2 Conceptual Analysis	10
1.2.3 Naturalism	12
1.3 Taxonomizing the Naturalism Debate	14
1.3.1 The Basic Schema	14
1.4 Situating the Competitors	16
1.4.1 Strong Naturalism	16
1.4.2 Non-Naturalism	17
1.4.3 Moderate Naturalism	19
Chapter 2: Kornblith	23
2.1 Introduction	24
2.2 The Exclusionary Argument Reconstructed	24
2.2.1 Knowledge as a Natural Phenomenon	26
2.2.2 The Purported Methodological Upshot	31
2.3 Attacking the Exclusionary Argument	35
2.4 Natural Kinds, Reductionism and the Status of Knowledge	36
2.4.1 Essentialism and Reductionism	37
2.4.2 The Multiple Realizability Objection	39
2.4.3 The Argument from Psychological Explanation	46
2.5 The Starting Points Objection	54
2.5.1 Kornblith, Conceptual Analysis and the Empirical Strategy	54
2.5.2 The Disanalogy Premise and Knowledge	56
2.5.3 Extramental Kinds, Starting Points and Relevancy	58
2.6 Conclusion	61
Chapter 3: BonJour	66
3.1 Introduction	66
3.2 BonJour (I): Against Naturalism	66
3.2.1 The Deficiency Argument	67
3.2.2 The Generality Subargument	69
3.3 BonJour (II): Rationalist Non-Naturalism	71
3.3.1 Platonism, Mental Content and Epistemic Access	71
3.3.2 A Priori Justification	75
3.4 Generality Again & the Methodological Upshot	76
3.4.1 Completing the Generality Subargument	76
3.4.2 The Purported Methodological Upshot	77
3.5 Attacking BonJour's Non-Naturalism	79
3.5.1 Setting Up	79
3.5.2 Against the Deficiency Argument	82
3.6 Conclusion	86

Chapter 4: Goldman	88
4.0 Review of the Negative Project	89
4.1 Introduction	89
4.1.1 Some Conceptual Presuppositions	91
4.2 Mentalism (I): an Overview	92
4.2.1 Conceptual Analysis: a Redux	92
4.2.2 Goldman's Basic Account	93
4.3 Mentalism (II): Arguing in its Favour	95
4.3.1 Evidentially Basic Beliefs	95
4.3.2 The Argument from Explanatory Comprehensiveness	97
4.3.3 The Argument from Parsimony	105
4.4 Objections and Replies	109
4.4.1 Objection # 1: "Isn't Mentalism too Subjective?"	109
4.4.2 Objection #2: The Reliability of Intuitions	111
4.5 Conclusion	113
Chapter 5: Conclusion	116
5.1 Review of the Basics	117
5.2 The First Prong (The Negative Project)	118
5.3 The Second Prong (The Positive Project)	120
5.4 Summarizing Remarks	121
Works Cited	122

Chapter 1: Background

1.1 Introduction

The contemporary debate over the compatibility of the a priori and conceptual analysis, on the one hand, and epistemological naturalism, on the other, features two broadly opposed camps or positions: those who take the a priori and conceptual analysis, on the one hand, and naturalized epistemology, on the other, to be mutually exclusive, and those who take the relevant doctrines to be reconcilable. And while this basic divide is relatively clear, it seems to me that much of the surrounding debate remains to be subject to systematic study. Thus, the following thesis will be an attempt to modestly contribute to the debate by proposing new a way of taxonomizing each camp and the family of more particular views which populate them. To this end, I shall subsequently call those who view the a priori and conceptual analysis as excluding naturalism, and vice versa ‘Incompatibilists’; and I shall call those who view the relevant doctrines as reconcilable ‘Compatibilists’. Ultimately, then, my hope is that the terms of the debate will become clearer for all interested parties once the metaepistemological terrain surrounding the Incompatibilist-Compatibilist divide is laid bare.

That being said, my aims in the following study will not simply be taxonomic or clarificatory. I will also argue in favour of a particular type of Compatibilism known as ‘Moderate Naturalism’. Developed in a series of papers by Alvin Goldman, Moderate Naturalism attempts to integrate the deliverances of the a priori and conceptual analysis into a naturalistically respectable account of epistemically basic sources of evidence. As such, I will make the case for a Goldman-style position in two steps, the former negative, and the latter positive. In the first step, my aim will be to create some logical space for Compatibilism by considering and rejecting two arguments for Incompatibilism courtesy

of Hilary Kornblith and Laurence Bonjour. In brief, both Kornblith and Bonjour argue in favour of an anti-Compatibilist principle which, as I shall maintain below, we have good reason to reject as inadequately motivated. In the second, my aim will be to exploit the logical space created in the first step by suggesting some positive reasons to think that a Goldman-style Moderate Naturalism is abductively superior to either of the Incompatibilist alternatives offered up by Kornblith or Bonjour. That is, I think we have good reasons, over and above the failure of Incompatibilism, to accept a Moderate Naturalism. To summarize my objectives, I will be advocating the following two-pronged thesis: (1) that both Kornblith and Bonjour's attempts to establish Incompatibilism fail, and (2) that Moderate Naturalism is abductively superior to its Incompatibilist rivals.

In terms of organization, my thesis shall proceed in the following manner. First, I will situate Kornblith, Bonjour and Goldman within a taxonomy centred around the Incompatibilist-Compatibilist distinction made above. The purpose of constructing this taxonomy will be clarificatory: it will be both necessary and helpful to compare and contrast the positions in question before considering my argument. Second, I will defend both prongs of my thesis at length; critically examining the arguments of Kornblith and Bonjour in turn before moving on to a discussion and defense of Goldman. Finally, in the third and final section I shall summarize the upshot and discuss what, in terms of the debate over naturalism in epistemology, I think can be gained from it.

1.2 Three Background Doctrines

I would like to begin by briefly sketching out the three doctrines which will structure the debate in the following chapters: the a priori, conceptual analysis and naturalism. This sketch will provide the requisite background through which both my taxonomy and argument shall be constructed.

1.2.1 The A Priori

Let us start with a general description of a prioriticity. The a priori is an *epistemic* notion, meaning, first and foremost, that it qualifies justification (in the knowledge-related sense), and second, that it should not be confused with necessity (which qualifies propositional truth-values).¹ So what we want is an account of what it is for a subject *S* to be justified in believing *p* a priori; where a prioriticity qualifies the *source* which yields *S*'s justified belief as output. With these considerations in mind, I offer the following basic definition of a prioriticity:

A proposition *p* is justified a priori for a subject *S* iff: i). *S* believes *p* with justification, and ii) empirical² experiences form no part of what makes *S* justified in believing that *p*.

This is to be distinguished from a posteriori justification, which can be defined as:

A proposition *p* is justified a posteriori for a subject *S* iff: i): *S* believes *p* with justification, and ii). empirical experiences form part of what makes *S* justified in believing that *p*.

¹ That is, following the Kripke-Putnam thesis, I reject the identification a prioriticity with necessity. Necessity is an explicitly *metaphysical* thesis; and although it often is coinstantiated with a priority it is not identical to the latter doctrine, which is explicitly *epistemic*. See Kripke, (1980) and Putnam (1975).

² Obviously, construing the a priori negatively in this manner requires a suitable definition of what exactly constitutes 'empirical experience'. Following Bonjour, I will stipulate that it includes any sort of process that is perceptual in the broad sense of being either: (a) A causally conditioned response to particular, contingent features of the world (b) Yielding doxastic states that have as their content putative information concerning such particular contingent features of the actual world. This allows us to rule out the possibility of conscious mental episodes per se. counting as experiential; they only count as such if they are 'perceptual' in the relevant sense. See Laurence Bonjour, *In Defense of Pure Reason*, Cambridge UK: Cambridge University Press, 1998, pg. 8.

As I alluded to above, this account makes essential reference to an *epistemic source condition*: minimally, a justified belief is a priori just in case it contains no empirical experiences (where ‘empirical experience’ is suitably defined) among its justifiers (i.e. among the sources of its justification). Defining the a priori by reference to source conditions has two distinct advantages: first, it does not presuppose that there *is* such a thing as a priori justification. Second, it does not specify what (assuming it should exist) a priori justification supervenes upon; it merely says what it does not supervene upon. This allows the definition to be useful to partisans of various competing epistemological theories (e.g., foundationalists and coherentists).

Assuming the existence of a class of beliefs which are justified a priori, a further question arises: what is the *target domain* of these beliefs? There are two broad views here which are salient for the proceeding discussion. The first version sees external reality as the domain of propositions over which a priori justification ranges (e.g. real properties, real modal relations, etc); whereas the second version sees the domain of propositions over which a priori justification ranges as internal reality (e.g. our own concepts and the relations between them).

1.2.2 Conceptual Analysis

According to what I will call (following Goldman) our ‘Standard Philosophical Methodology’, or SPM³, a central aim of philosophical theorizing is the production of plausible “conceptual analyses” of target concepts of philosophical interest, i.e., accounts of the necessary and/or sufficient conditions under which such concepts properly apply

³ See Alvin Goldman and Joel Pust, ‘Philosophical Theory and Intuitional Evidence’. In M. DePaul & W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield, 1998, pp. 180-183.

which are not only illuminating but also justified. And, also according to SPM, the means whereby we produce such conceptual analyses, i.e., the method of doing conceptual analysis, inherently involves reliance on a priori justified beliefs in the form of “application intuitions”—intuitive judgments about the application (or non-application) of the relevant concepts in particular, hypothetical and actual scenarios.⁴ First, a scenario e is considered or constructed and application intuitions regarding the application (or non-application) of the theorist’s concept F to scenario e are elicited.⁵ Second, the intuitions are then taken, inferentially, as evidence for the soundness of her explicit theory or analysis of F ; where a negative application intuition in e constitutes *prima facie* evidence against F , should the latter suggest positive application in e . In other words, application intuitions are presented in SPM as basic, a priori sources which can lead (inferentially) to a priori justified beliefs in the truth of the conceptual analyses they cohere with, and to a priori justified beliefs in the falsity of the conceptual analyses they conflict with.

1.2.3 Naturalism

While epistemological naturalism resists easy summarization, it can be characterized, I think, as containing two core theses: one metaphysical, the other methodological.

The metaphysical thesis, briefly, states that the only entities (e.g. objects, properties, relations, events, states of affairs, etc.) that epistemological theories should

⁴ Special thanks to David Matheson for helping me to clarify the relationship between application intuitions and conceptual analyses in SPM.

⁵ In the fourth chapter, I shall have more to say about how an application intuition can be said to express or reflect the content of the concept being deployed. For now, we can roughly define application intuition as a species of occurrent or spontaneous judgement which expresses the semantic content of the relevant concept underpinning its deployment.

quantify over (i.e., count as real) are those countenanced by the natural sciences and whatever supervenes upon them⁶. On the assumption that the natural sciences countenance only physical entities,⁷ metaphysical naturalism can thus broadly be taken as implying a commitment to an ontological physicalism of the kind operative in the natural sciences:

(Metaphysical Thesis): Epistemological theories should only quantify over the physical entities countenanced by the natural sciences and whatever supervenes on these entities.

Moving on, the methodological thesis states, roughly, that epistemology should adopt the largely empirical⁸ methods characterizing inquiry in the natural sciences. This thesis naturally follows from commitment to ontological physicalism, insofar as the structure of the relevant target domain (comprised of largely physical entities) demands a particular set of approaches optimized for this target domain. In abbreviated form:

(Methodological Thesis): Epistemological methods should involve the (largely empirical) methods of the natural sciences.⁹

One may note here that my characterization omits a precise specification of what the relevant empirical methods are. I shall not, however, have the space to provide the relevant specification (despite its manifest importance); and as such the basic account given here will have to suffice.

⁶ This will be discussed further below. I adopt a general framework for naturalism which is on a par with Papineau (1993) and Goldman (1999).

⁷ See Hilary Kornblith, *Naturalism, Both Metaphysical and Epistemological*, Midwest Studies in Philosophy 19 (1), 1994, pp. 40-42.

⁸ Where 'empirical' should be read in an inclusive manner, including methods suited to the study of broadly physical and supervenient entities. I realize that this is vague, but much of the literature surrounding naturalised epistemology is structured in a similar manner, and I don't have the space here for a drawn out discussion.

⁹ For a similar view and more in-depth survey of the core theses of epistemological naturalism, see Kornblith (1994), Devitt (1980) and Papineau (1993), (2009).

1.3 Taxonomizing the Naturalism Debate

With the basic doctrines I will be concerned with sketched out, I will now commence with the classificatory element of my argument. The task here will be to schematize the various positions in the naturalism debate and discuss the theoretical commitments they assume.

1.3.1 The Basic Schema

I identify two basic positions or categories in the debate: Incompatibilism, which argues for the mutual exclusion of a priori justification and conceptual analysis (hereafter, I shall use ‘a priori analysis’ to designate the combination of the two doctrines), on the one hand, and naturalism on the other, and Compatibilism, which rejects mutual exclusion and instead endorses the compatibility of the three relevant doctrines. As these two categories are very broad, they can be further broken down into individual metaepistemologies as a function of their particular theoretical commitments. As such, I identify two types of Incompatibilism, a naturalistic version, exemplified by the work of Hillary Kornblith, and a non-naturalist version, exemplified by the work of Laurence Bonjour. I only consider one Compatibilist position here, the Moderate Naturalism of Alvin Goldman, but this should not be taken as implying that it is the only possible Compatibilist option. My proposed taxonomy can be clarified through the following table:

- (i) Naturalism (a commitment to the Metaphysical Thesis and to the Methodological Thesis).
- (ii) A Prioricity (a commitment to a priori justification).

- (iii) Conceptual Analysis (a commitment to the viability of conceptual analysis as it appears in SPM).

Incompatibilism

(Kornblith) Strong Naturalism – Deny (ii)-(iii), Affirm (i).

(BonJour) Non-Naturalism – Deny (i), Affirm (ii)-(iii)

Compatibilism

(Goldman) Moderate Naturalism – Affirm (i)-(iii)

Let me now make a couple of important points. First, note that Incompatibilism unites two otherwise opposing views as unlikely allies against Compatibilism (Moderate Naturalism): Kornblith and BonJour both embrace Incompatibilism, despite having otherwise widely differing theoretical commitments – the former rejects a priori analysis on naturalistic grounds while the latter rejects naturalism on broadly Rationalist or Platonistic grounds.

Second, one may also wonder how Incompatibilism is motivated. As I have mentioned, it is based upon the assumption that (ii)-(iii) and (i) are mutually exclusive; i.e. committing to (i) entails a rejection of (ii)-(iii) and vice versa. In other words, I believe that Incompatibilists would be committed to something like the following anti-Compatibilist principle:

(ICA): Epistemological naturalism (i) and a priori analysis (ii)-(iii) are mutually exclusive: the adoption of naturalism rationally requires the rejection of a priori analysis and vice versa.

What about Compatibilism? Turning now to Goldman, we can note that a position like his would reject commitment to a principle approximating (ICA). In other words, a Goldman-style Moderate Naturalism (of the type I will be defending) takes the position that there is no (at least *prima facie*) incompatibility between doctrines (i) and (ii)-(iii).

Now obviously, given the opposite claims being made here, it can't be that Incompatibilism and Compatibilism are somehow both true; and the purpose of my taxonomy is thus to lay the conflict bare and provide a springboard for my defense of Compatibilism (i.e. Moderate Naturalism) and the rejection of (*ICA*). However, I think I have here made the topography of the debate clear enough, and we can now take a closer look at each individual position.

1.4 Situating the Competitors

The final element of this chapter will be a basic survey of the core arguments standing behind each metaepistemological position in consideration here. This survey will provide background for the line of attack I will be concerned with in the rest of the thesis.

1.4.1 Strong Naturalism

I take the work of Hilary Kornblith as representative of Strong Naturalism; where this entails the rejection of (ii)-(iii), i.e., a priori analysis, on grounds of incompatibility with (i).

Kornblith argues that the relevant target domain of a priori analysis (i.e., the domain of entities which application intuitions, and the conceptual analyses they support, are about) is populated by empirical *natural kinds*. Now importantly, he takes this natural kinds view to place significant adequacy constraints upon any naturalistically respectable metaepistemology for the following reason. Consulting our a priori application intuitions regarding a predicate can *only* generate and justify inferences ranging over what is conceptually possible. However (assuming that knowledge is in

fact a natural kind), Kornblith argues that what we are *really* interested is an a posteriori and empirical account of the phenomenon as it exists here “in the world”.¹⁰ Thus – and this is the crux of his position – Kornblith *denies* that the deliverances of a priori analysis can be justified¹¹: while we will have application intuitions about the extension of a predicate at some scenario *e*, they carry no evidential weight because, to continue the argument, all justified beliefs about a particular concept *F*, where *F* is a natural kind, must ultimately derive from empirical sources.

Importantly, then, on Kornblith’s view commitment to a robust naturalism *entails* a rejection of a priori analysis: because the targets of inquiry are natural kinds, a priori analysis has no place in epistemology proper. I call this line of reasoning the ‘Exclusionary Argument’¹²:

- (1). If epistemology is to be methodologically naturalized in a thoroughgoing manner, then epistemology cannot retain scientifically suspect methods
- (2). Epistemology is to be methodologically naturalized in a thoroughgoing manner
- (3). Therefore, epistemology cannot retain scientifically suspect methods.
- (4). If an epistemological method cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain, then it is a scientifically suspect method.
- (5). *A priori* analysis cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain.
- (6). Therefore, a priori analysis is a scientifically suspect method.
- (7) Therefore, epistemology cannot retain a priori analysis.

¹⁰ Hilary Kornblith, *Knowledge And Its Place in Nature*, Oxford UK: Oxford University Press, 2002, pp. 13-20.

¹¹ Note that Kornblith doesn’t deny that we have a priori beliefs per se.; rather his thesis is that they carry no evidential weight whatsoever. This is an important difference between Kornblith and ‘radical empiricist’ proponents of naturalized epistemology such as Quine (at least Quine on a popular reading) who seem to propose a complete elimination of the a priori. See BonJour (1998), chapter 3 and Quine, (1969), chapter 3.

¹² Thanks to David Matheson for some useful suggestions about the formulation of this argument.

Kornblith rejects (ii)-(iii) as a result of the Exclusionary Argument: construing epistemic kinds (like knowledge) as natural phenomena leaves no room for the non-empirical methodology of a priori analysis.

1.4.2 Non-Naturalism

Laurence Bonjour is a longstanding proponent of what I call epistemological Non-Naturalism. Non-Naturalism affirms Incompatibilism about naturalism and traditional methods, but draws the opposite moral to Strong Naturalism: given the essentiality of a priori analysis to a fundamentally non-skeptical epistemology, (i) is to be rejected and (ii)-(iii) retained.

BonJour makes the case for this conclusion negatively through an argument from elimination. He maintains that because radical empiricism (which rejects the a priori *simpliciter*) leads to radical scepticism and moderate empiricism (which rejects a rationalist view of the a priori¹³) untenably collapses into either a radical empiricism or a concealed version of rationalism,¹⁴ no form of naturalism can provide a genuinely viable metaepistemology. Thus, concludes Bonjour, we are forced to concede that Non-Naturalism is the only viable and non-sceptical metaepistemological position to take.

The key premise motivating this argument for Non-Naturalism is Bonjour's Platonism in which the target domain of application intuitions is an abstract, non-

¹³ I.e. a moderate empiricism, according to Bonjour, will be such that it limits the scope of a priori justification to so-called 'analytic' propositions; that is, propositions which are putatively justified by virtue of logical form ("all bachelors are unmarried males"). The point of this was to retain a type of a priority which didn't make any claim to elucidate the structure of external reality (proponents of moderate empiricism taking this to be a dubious thesis).

¹⁴ See Bonjour, 1998, pp. 58-61.

spatiotemporal realm populated by universals which, through the evidential route provided by spontaneous “intellectual apprehension” or “rational insights” – application intuitions as Bonjour thinks of them, act as the justifiers for a priori beliefs.¹⁵ On Bonjour’s view, then, given the failure of naturalism we are forced (by elimination) to adopt a Non-Naturalist conception of (ii) and (iii)¹⁶ (where the instantiation of abstract universals allows our metaepistemology to avoid scepticism). His argument for this position I dub the ‘Deficiency Argument’:

- (1). Either (Radical) Naturalism is a minimally adequate meta-epistemological theory or Non-Naturalism is true.
- (2). If Naturalism is to be minimally adequate, then it mustn’t imply scepticism
- (3). Naturalism implies scepticism.
- (4). Therefore Naturalism isn’t minimally adequate.
- (5). Therefore Non-Naturalism is true.

In sum, Bonjour embraces Incompatibilism on the strength of his contention that naturalism implies scepticism; for it follows from this premise that any non-sceptical metaepistemology will entail non-naturalism.¹⁷

1.4.3 Moderate Naturalism

Third is the position which I wish to defend, i.e. Moderate Naturalism. Inspired primarily by the work of Alvin Goldman, Moderate Naturalism differs from the two

¹⁵ Ibid, pp. 180-185.

¹⁶ Note that while Bonjour’s direct target isn’t naturalized epistemology per se., but rather “empiricism” (broadly construed), it is quite clear that he sees the two doctrines as mutually entailing. For as Bonjour depicts the dialectical situation, moderate and radical empiricism are both motivated primarily by a desire to account for the a priori in a naturalistically acceptable way, one which therein limits the scientifically suspect elements of rationalism. See Bonjour, 1998, chapters 2 and 3.

¹⁷ Ibid, chapter 4., especially pp. 175-181.

previously considered positions by endorsing Compatibilism: we can conjointly *affirm* the set of (i)-(iii) and endorse a naturalistically respectable account of a priori analysis.

As I mentioned above, my argument for Compatibilism will have two components; one negative, where I will maintain that we have good reason to reject the (ICA) principle, and one positive, where I will describe my preferred alternative in detail and argue that it has distinct advantages. Let me describe each component in more detail:

Moderate Naturalism (1): The Negative Project

In accordance with the (ICA) principle, Kornblith and Bonjour both give arguments designed to show that (i), (ii) and (iii) form an inconsistent triad.

Kornblith's Exclusionary Argument for Strong Naturalism maintains that a priori analysis entails Non-Naturalism insofar as it falls outside of the scope of practices which yield justified beliefs about natural kinds. I will object to this argument by maintaining that its premise (5) is false: it presupposes that the targets of epistemological theorizing (i.e. epistemic kinds) are, in fact, natural kinds, and I will argue that we *do not* have good reason to think that knowledge and other epistemic kinds are natural kinds. Thus, because premise (5) fails, Kornblith's motivation for viewing a priori analysis as naturalistically suspect falls away, allowing us to reject it wholesale.

BonJour attempts to establish the entailment in question via the Deficiency argument. More specifically, he provides an elimination argument to the effect that *only if* we accept a Non-Naturalist version of a priori analysis can we defeat scepticism. My objection here shall rest upon demonstrating the failure of a subargument invoked by

BonJour in support of the key premise (3) underpinning the Deficiency Argument. That is, I will argue that BonJour isn't successful in demonstrating the deficiency of naturalism *per se*.

My overall aim here is thus to demonstrate that we have no compelling reason from either Kornblith or BonJour, at the end of the day, to accept the (*ICA*) principle.

Moderate Naturalism (2): The Positive Project

The second component of my argument will then build on the logical space opened up by the failure of Incompatibilism to make the case for a Moderate Naturalist harmonization of (i)-(iii)), and thus to provide some positive reason to reject (*ICA*).

My starting point is a conception of a priori justification which (following Goldman) sees the relevant target domain as the internal reality of concepts and relations. This view, which I call 'Mentalism', claims – in contrast to both Strong Naturalism and Non-Naturalism – that the target domain of application intuitions and the conceptual analyses that are used to support them is explicitly *mental* (psychological), thus invoking no extra mental abstracta or natural kinds¹⁸ So construed, I argue that Moderate Naturalism can thus provide a straightforward explanation of how a consistent conjunction of (i)-(iii) is possible: we can admit (ii)-(iii) into our metaepistemology *without* violating the constraints of (i) because Mentalism, acting as the crucial bridging doctrine, provides the resources for a *naturalistically acceptable* explanation of how an a priori evidential route from application intuitions about a concept *F* to plausible analyses of *F* itself is both possible and feasible.¹⁹ The Mentalist account of the relevant

¹⁸ See Goldman and Pust, 1998, pp. 188-190.

¹⁹ Ibid, pp. 188-189.

justificatory process, in other words, provides a straightforward “in the mind” explanation of the evidential route from evidence to analyses of a concept of philosophical interest.²⁰

More specifically, I will first contend that Mentalism – as the account of a priori analysis underpinning Moderate Naturalism – is more explanatorily comprehensive than either Kornblith or Bonjour’s Extra-Mentalism. Second, I will contend that Mentalism is also more ontologically parsimonious. Overall, then, I will maintain that we have good reason to hold that Moderate Naturalism (i.e. Mentalism,) is *abductively superior* to its rivals.

²⁰ A worry might be voiced here to the effect that because Mentalism construes the target domain of a priori beliefs as mental concepts, it therein collapses into an epistemic anti-realism given that it construes the targets of conceptual analysis as mental phenomena. Although I won’t have time to discuss these complex issues in this thesis, note that I don’t believe that Mentalism collapses into anti-realism, at least as long as we adopt an externalist and dispositionalist account of epistemic justification.

Chapter 2: Kornblith

2.1 Introduction

The first Incompatibilist position that will be considered is Kornblith's Strong Naturalism. To recap: Kornblith, via the Exclusionary Argument, maintains that a consistent naturalism (i.e. (i)) entails the exclusion of a priori conceptual analysis (i.e. (ii)-(iii)), on grounds that adoption of the latter doctrines commit one to naturalistically suspect methods of inquiry.

My primary aim in this chapter will be to show why this Exclusionary Argument *fails* to support Kornblith's assumption of Incompatibilism (i.e. his commitment to the (ICA) principle), thus removing a substantial naturalistically-motivated barrier to Compatibilism. The argument will proceed in two stages: first, (2.2) I will examine Kornblith's defence of Incompatibilism in some detail. Second, I will then proceed (2.3-2.4) with arguments designed to show that it does not successfully establish what Kornblith needs in order to set his Strong Naturalism on firm ground.

2.2 The Exclusionary Argument Reconstructed

As we have seen, Kornblith thinks that a priori analysis is naturalistically suspect and must be discarded because it cannot yield justified beliefs about epistemic kinds considered as natural kinds. The Exclusionary Argument attempts to establish this by moving from (a) a *metaphysical* premise about the status of knowledge as a natural kind to (b) a *methodological* conclusion stipulating adequacy constraints upon the proper domain of naturalistic inquiry. Put briefly, Kornblith argues that because a priori

analysis cannot justify beliefs about natural kinds, it therefore falls outside of the naturalist's (scientific²¹) methodological purview and must be discarded.

To sharpen understanding of this point, consider the Exclusionary Argument as introduced in Chapter 1:

- (1). If epistemology is to be methodologically naturalized in a thoroughgoing manner, then it cannot retain scientifically suspect methods
- (2). Epistemology is to be methodologically naturalized in a thoroughgoing manner
- (3). Therefore, epistemology cannot retain scientifically suspect methods.
- (4). If an epistemological method cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain, then it is a scientifically suspect method.
- (5). *A priori* analysis cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain.
- (6). Therefore, *a priori* analysis is a scientifically suspect method.
- (7) Therefore, epistemology cannot retain *a priori* analysis.

Reconstructing the Exclusionary Argument in this manner allows us to see exactly how Kornblith infers methodological adequacy constraints from explicitly metaphysical considerations. The key to the inference in question is premise (5): notice that it presupposes that target terms of epistemological interest, such as 'knowledge', 'knows', 'justified belief' etc. refer to natural kinds. In affirming the antecedent of the conditional statement made in (4) on explicitly metaphysical grounds (i.e. knowledge as a natural kind), then, (5) sets up Kornblith's overall conclusion that a methodologically naturalized epistemology cannot retain *a priori* analysis. In other words, from a kind identification

²¹Where 'scientific' should be taken, broadly speaking, as comporting with my discussion of Naturalism in the first chapter. See David Papineau, "Naturalism", *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/naturalism/>.

comes a methodological constraint upon the naturalization project: a priori methods are fundamentally unsuited to the study of a natural, empirical subject matter and should be dropped by the naturalist.

Given, then, the importance of the epistemic kinds as natural kinds view presupposed by (5) in the Exclusionary Argument, Kornblith attempts to independently motivate the premise with a subargument motivated by evolutionary considerations. I now want to consider his subargument in some detail, for coming to grasp with it will be of great importance to my critique of the Exclusionary Argument below.

2.2.1 Knowledge as a Natural Phenomenon

Kornblith, as I read him, advocates the following theory of knowledge: it is a reliable, information-bearing cognitive state that is the product of natural selection; a state that, in playing an explanatory role in the prediction of animal behavior, is best viewed as a natural kind.

The argument has three stages. First, Kornblith stipulates that Darwinian natural selection is a plausible theoretical explanation for understanding how organisms possess truth-conducive informational states (or ‘beliefs’, in Kornblith’s terminology) about their external environment. Second, building upon this evolutionary view, he argues that these selected-for beliefs are, in fact, best interpreted as a type of species-level knowledge. Third, Kornblith concludes that knowledge is properly viewed as a *natural kind*: it both plays an integral and a unifying role in the prediction and explanation of the successful adaptation of organisms to environmental requirements.

Natural Selection

Natural environments make substantive informational demands upon the organisms which inhabit them, demands which require that successful organisms process external inputs, integrate these inputs as beliefs, and organize behavior accordingly; where this includes (at minimum) survival and reproduction.²² Given that many organisms are successful in meeting these demands, what would adequately explain this capacity?

Kornblith invokes Darwinian natural selection:

As Quine suggests, "There is some Encouragement in Darwin." Creatures whose belief generating mechanisms do not afford cognitive contact with the world have a pathetic but praiseworthy tendency to die before reproducing this kind. Since believing truths has survival value, the survival of the fittest guarantees that our innate intellectual endowment gives us a predisposition for believing truths.²³

As suggested here, the rough idea behind Darwinian natural selection is that successful organisms have evolved over tens of thousands of years to possess biological or cognitive processes which bias or have a predisposition towards generating *true* beliefs about their external environment. The reason is that true beliefs about the external world allow an organism to negotiate its environment and reproduce successfully (where, generally, false beliefs wouldn't). So, the positive survival value of these processes is such that they would have been selected for at the species level via the process of evolutionary adaptation.²⁴

Darwin's theory of natural selection thus provides a powerful tool for explaining the tendency for organisms to possess processes which (in general) yield mostly true

²² Roughly, the proper scope of Darwinian evolution is the level of species, rather than individuals. That is, while individual genetic variation is the primary means by which evolutionary pressures are exerted, this genetic variation produces or manifests itself as group or species-level traits. For more, see Sober and Wilson (1998).

²³ Hilary Kornblith (Ed.), *'Epistemology Naturalized'*, Cambridge, MA: MIT Press, 1985, pg. 4f.

²⁴ Hilary Kornblith, *Knowledge And Its Place In Nature*, Oxford UK: Oxford University Press, pp. 56-61.

beliefs about their environment – producing true beliefs is conducive to survival and reproduction.²⁵

Knowledge as External Informational State

The second stage turns on a further, distinctly epistemological question: assuming that organisms display an evolved capacity to form true beliefs about their external environments, can we classify this capacity as a kind of knowledge? Kornblith answers in the affirmative. I begin with a revealing passage:

...Explanation of successful behavior... requires the notion of knowledge rather than mere belief... [The] true beliefs that particular [animals] acquire will be the product of a stable capacity for the production of true beliefs. The resulting true beliefs are not merely accidentally true; they are produced by a cognitive capacity that is attuned to its environment. In a word, the beliefs are reliably produced. The concept of knowledge which is of interest here *thus requires reliably produced true belief...*²⁶

For Kornblith, the property which converts true beliefs (considered as informational states) into knowledge is *reliability*. The argument here is based upon explanatory ascent. When we consider the cognitive capacities of token organisms, the epistemic category “mere belief” is explanatorily adequate. This is because the cognitive abilities of an individual organism vis-a-vis its particular environment are sufficiently “accidental” (presumably through the workings of genetic variation) to be captured by a term compatible with epistemic luck. When we consider the cognitive capacities of a species, however, the category “mere belief” becomes explanatorily deficient. This is because the cognitive capacities a successful species displays are, when considered *as a*

²⁵ Ibid, pp 56-59.

²⁶ Ibid, pp. 57-58. Italics added for emphasis.

*type*²⁷, sufficiently non-accidental and reliable to confer a higher degree of epistemic value upon the informational states they produce.

Note that Kornblith's conception of epistemic justification – of whatever it is beyond true belief that yields knowledge – and hence his conception of knowledge, is firmly externalist: as long as an internal belief-forming process meets the external condition of reliability, the true beliefs it produces will be justified and amount to knowledge.²⁸

External Informational States as Natural Kinds

The third phase of Kornblith's argument makes the case for knowledge as a natural kind. It runs as follows. First, two necessary conditions for natural kind membership based on a "cluster kind" theory are stipulated. Second, he argues that 'knowledge' denotes a kind which satisfies the two aforementioned conditions.

Beginning with the initial stipulation of conditions for natural kindhood, Kornblith says the following. A predicate denotes a natural kind if and only if two conditions are satisfied: (i) the predicate refers to a set of robustly co-occurring and mutually supporting properties which form a "homeostatic cluster"²⁹; and (ii), these co-occurring properties are "projectable"; i.e. they persist through external change and figure centrally in causal/nomological explanations. In other words, then, a natural kind

²⁷ Ibid, pp. 61-62.

²⁸ See *Knowledge and its Place in Nature*, pp. 61-69.

²⁹ Cluster Kinds theory attempts to reconcile the semantic intuitions implied by causal theories of reference with biological arguments for kind-membership indeterminacy. More specifically, proponents such as Richard Boyd (1982, 1988) argue that while a particular kind *t* does have an underlying physical structure, membership of *t* is determined by the presence of the relevant set of clustered properties. The insight here is that there are no strict necessary and sufficient conditions for membership of *t*; rather it is the presence of a property cluster approximating other members of the set. Kinds, then, are a posteriori but vague. See *Knowledge And Its Place In Nature*, pp. 61-62.

is an entity constituted by set of “bundled” properties which are both mutually supporting, causally efficacious and demonstrate a regularity which supports inductive inference.

Relying on this basic cluster kinds framework, Kornblith argues that the class of informational states designated by the process-related state type ‘Knowledge’ displays all of the relevant characteristics of causally efficacious homeostatic clusters, and thus can be properly designated a natural kind. He says:

Animal knowledge... [is] such a... category, a category that features prominently in causal explanations, and thus in successful inductive predictions. If we wish to explain why it is that members of a species have survived, we need to appeal to the causal role of the animals’ knowledge of their environment... The knowledge that members of a species embody is the locus of a homeostatic cluster of properties: true beliefs that are reliably produced, that are instrumental on the prediction of behavior successful in meeting biological needs and thereby implicated in the Darwinian explanation of the selective retention of traits.³⁰

And furthermore:

Understanding what that theoretical unity is is the object of our study, and it is to be found by careful examination of the phenomenon, that is, something outside of us, not our concept of the phenomenon, something inside of us. In short, I see...knowledge... [as a] natural kind.³¹

In other words, the process-type ‘Knowledge’ denotes an external (in the epistemological sense outlined above) set of co-occurring and mutually supporting homeostatic properties, a set which features prominently in the prediction and causal explanation of the behavior of complex organisms.

Thus, according to Kornblith, knowledge – the kind that stands as the referential content of ‘knowledge’ – satisfies the conditions required to be a natural kind according to his preferred account.

³⁰ Ibid, pg. 62.

³¹ Ibid, pg. 11. Italics added for emphasis.

2.2.2 The Purported Methodological Upshot

With the three-part subargument for knowledge as a natural kind outlined in 2.2.1, we can now see more clearly what work it does in providing justificatory support for the crucial premise (i.e. (5)) of the Exclusionary Argument, and consequently how Kornblith uses it to draw substantive methodological constraints upon the naturalization project.

Kornblith's inference from knowledge as a natural kind to a methodological sanction of a priori analysis is, on the face of it, relatively straightforward: establishing that knowledge is a cluster-kind of natural properties implies that we *can't* learn anything about it via a priori analysis. This is because such cluster-kinds are individuated by local empirical conditions; meaning that they can only be discovered via a posteriori investigation of the external (i.e. actual) world. As a result, analyzing our pre-empirical concept of knowledge in isolation of such an investigation cannot yield any justified deliverances about the target phenomenon:

..Since our ultimate target is the phenomenon [itself]...we would do better to study those extra-mental phenomena directly... The investigator who is interested in aluminum will learn little about his target by studying folk concepts of aluminum. He will learn more by studying the concepts of sophisticated chemists... [and] to look at aluminum itself.³²

Just as we must investigate the *kind* Aluminum rather than the concept, we must investigate the *kind* Knowledge, not the concept; for concepts cannot illuminate natural kinds, only the world can.

So, the subargument for knowledge as a natural cluster-kind outlined in 2.2.1 (c), i.e.:

³² Hilary Kornblith. *Naturalism and Intuitions*. In C. Beyer & A. Burri (eds.), *Philosophical Knowledge: Its Possibility and Scope*. Grazer Philosophische Studien, Amsterdam: Rodopi, 2007, pg. 36.

(A) If a term denotes a co-occurring set of properties which: (i) constitute a “Homeostatic Cluster”, (ii) this set is “projectable”; i.e. it persists through external change and figures centrally in causal/nomological explanations, then this term denotes a natural kind.

(B) The term ‘knowledge’ denotes a co-occurring set of properties which satisfy conditions (i) and (ii).

(C) Therefore, ‘knowledge’ denotes a natural kind.

appears to be pulling the justificatory weight in support of Kornblith’s methodological stricture contained in premise (5) of the exclusionary argument:

(5). *A priori* analysis cannot illuminate the natural kind(s) that stands as the referential content of the term(s) contained within its target domain.

This can’t be the full extent of Kornblith’s argument, however. For the cluster-kinds view argued for in (a)-(d) is – and this point I stress – without further supplementation, *too weak* to support the methodological stricture expressed in premise (5). The reason is that cluster-kinds *per se.* are conceptually compatible with a token physicalism which construes the relevant properties in higher-order (descriptivist) and functional terms; in which case implying that they would be amenable to a priori analysis. More specifically, the reason why cluster-kinds construed in higher-order or functional terms would fail to rule out the relevant methodological stricture is based upon semantic considerations. For because (assuming token physicalism) such higher-order properties are not extensionally equivalent to the base properties which instantiate them, the reference of their associated terms is *mediated* (in the Russellian descriptivist sense) by their descriptive (intensional) content. This is very important, for it implies that propositions regarding the extension of the particular higher-order term can be justified *a priori* by analysing the relevant descriptive content of the scenario in question; blocking Kornblith’s desired conclusion.³³

³³ For more, see Kripke (1980).

Subsequently, on further examination it becomes clear that Kornblith intends to strengthen the cluster kind view of knowledge by committing to a causal theory of reference and essentialism about natural kinds. The following passage, I think, makes this quite explicit:

[Conceptual analysis] is just one way to mark [the extension of a term]; causal theorists of reference have another (and to my mind, better) account... it remains to be shown why our folk epistemological notions are of epistemological interest in their own right. Why should our folk epistemological notions be of any more interest to epistemologists than our folk chemical notions are to chemists?³⁴

Adopting the semantics of a causal theory of reference strongly suggests that Kornblith considers each property cluster-type to be *extensionally equivalent* (type-identical) to an essential physical or compositional kind (viz. 'kind essentialism').³⁵ Roughly, the Causal theory of reference claims that proper names and natural kind terms are semantic primitives, meaning that they directly pick out or denote their bearer. This is in contrast to Descriptivist theories of reference, which claim that such terms denote via the satisfaction of a semantically complex description. So, on the Causal account, a natural kind term denotes what it does not through the satisfaction of some complex description "in the head", but rather through a causal chain leading back to an initial baptism or designation. Now, importantly, because natural kind terms do not denote via description, a natural kind term can refer to one natural kind in one world and a different natural kind in another world even if the kinds referred to have exactly the same descriptive properties

³⁴ Kornblith, *Knowledge And Its Place In Nature*, pp. 18-19. Also, consider the following suggestive passage from his *Replies to Alvin Goldman, Martin Kusch and William Talbott*, pg. 438: "Just as the sciences empirically discover the essential properties of natural kinds such as water, an empirical investigation of the phenomenon of knowledge may discover the very features which make it what it is. On my view, a proper account of knowledge would give us an account of its necessary properties, just as a proper scientific account of the nature of water properly identifies it as H₂O. And just as water is a natural kind in every world if it is a natural kind in this world, knowledge is a natural kind in every world if it is a natural kind in this world." For more see *Replies to Alvin Goldman, Martin Kusch and William Talbott*, *Philosophy and Phenomenological Research*, Vol LXXI, no. 2, 2005.

³⁵ More on this below but see Kripke (1980), Putnam, (1975), (1983).

or ‘modes of presentation’ (e.g. ‘Clear, potable, drinkable liquid’...); the kind referred to in each case is instead individuated by whatever *compositional structure* is picked out in the particular modal scenario in question (‘H2O’ or ‘XYZ’). This is why the Causal theory is essentialist: all and only those worlds containing H2O are worlds in which ‘water’ refers to water, for water just *is* H2O – ‘water’ is a “rigid designator”. This move, I emphasize, is the key to Kornblith’s defence of the purported methodological upshot of the natural kinds view of knowledge; for it provides the strong modal tie between property clusters and compositional kinds that he needs. More specifically, it allows Kornblith to claim that the meaning of ‘Knowledge’, as an unmediated semantic primitive, is a direct function of its referential content (i.e. the constituent of the proposition is just the object itself); which in turn is determined by the local compositional kind the term rigidly designates.

It follows from this account, then, that ‘knowledge’ has as its referential content a phenomenon that is only capable of being investigated a posteriori; for we only, epistemically speaking, come to know the relevant essential kind (and thus the referential content of the term) by *empirical investigation of the actual world*.

As a result, adopting the conjunction of the causal theory of reference and essentialism provides the bridge between knowledge as a natural kind and premise (5), i.e. the methodological stricture against a priori analysis: because the meaning of natural kind terms is determined by its essential, actual-world structure, there is no way for our a priori semantic intuitions to illuminate the target domain in any informative way. That is, our a priori application intuitions about knowledge are unreliable in the same way that

our a priori application intuitions about chemical kinds are unreliable: the target phenomenon is fundamentally empirical and a posteriori.³⁶

Thus, instead of proceeding from (A)-(C) alone (as it appears *prima facie*), there are two further premises in Kornblith's subargument:

(D) The referential content of natural kind terms is discovered empirically by investigating their locally-determined essence [causal semantics + kind essentialism].

(E) If the referential content of natural kind terms is discovered empirically by investigating their locally-determined essence, then the referential content of such terms *cannot* be illuminated by a priori analysis.

From which Kornblith derives the key premise of the Exclusionary Argument (hereafter I shall designate (A)-(E) the 'Natural Kinds' subargument):

(5). A priori analysis cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain.

2.3 Attacking the Exclusionary Argument

With the Natural Kinds subargument supporting premise (5) of Kornblith's Exclusionary Argument laid out, I will now contend that we have good reason to reject it. Rejecting this subargument is significant because premise (5), as I indicated above, provides key motivation for the methodological sanction Kornblith advocates. Thus, undermining the Natural Kinds account of knowledge will allow me to challenge the Exclusionary Argument and reject its purported methodological upshot.

³⁶More accurately, Kornblith denies that a priori conceptual analyses can be justified on the basis of application intuitions. So, he doesn't deny that we *can* have application intuitions about the extension of a concept/predicate *F* at some scenario *e*; what he is denying is that they carry any evidential weight because (following from knowledge as a natural kind) all justified beliefs about a particular concept *F*, where *F* denotes a natural kind, must ultimately derive from empirical sources.

My objection to the Natural Kinds account of knowledge will employ two broad lines of argumentation. First (2.4), I will grant Kornblith's position and maintain that once we tease out its strongly reductionist consequences, the Natural Kinds view is revealed to be a highly implausible theory of knowledge (on grounds of explanatory adequacy). Second (2.5), I shift gears somewhat, and argue that Kornblith implicitly relies upon a priori analysis. This is significant because it implies that Kornblith is no more able to justify his own substantive theory of knowledge (vis-a-vis a fundamentally extramental target phenomenon) than his supposed competitors.

2.4 Natural Kinds, Reductionism and the Status of Knowledge

The initial line of argumentation I will bring to bear against Kornblith is broadly abductive in character, and is designed to draw out the highly implausible consequences of his epistemological views. More specifically, I will show that Kornblith's Natural Kinds account (through its commitment to essentialism) is a form of reductionism about knowledge, and contend that such a reductionism with regards to knowledge is theoretically implausible.

This implausibility will, in turn, be drawn out in two separate ways. After discussing essentialism and reductionism (2.4.1), I will put forth two anti-reductionist arguments adopted from similar debates in the philosophy of mind.³⁷ First (2.4.2) I will argue, via multiple realizeability, that Kornblith's reductionism severely constrains the potential equivalence-class of knowers to all and only those entities which share an identical compositional kind; and furthermore that this theoretical constraint is both

³⁷ By similar debates, I mean similar anti-reductionist (anti type-physicalist arguments) in the philosophy of mind.

arbitrary and most likely false. Second (2.4.3), I argue that reductionism about knowledge has the further undesirable consequence that it arbitrarily stops us from making important theoretical generalizations about the epistemic property denoted by 'Knowledge'.

2.4.1 Essentialism and Reductionism

As I argued in the previous section, Kornblith commits himself to the semantics of the causal theory of reference and kind essentialism about natural kinds in order to support the methodological sanction against a priori conceptual analysis that he advocates. And although it is intuitive that such a kind essentialism amounts to a form of reductionism, the link between these doctrines is generally left implicit in the relevant literature. It is important for my following objections, however, so let me pause and discuss it.³⁸

A good way to flesh out the relationship is to consider the modal status of primary and secondary nomic properties.³⁹ Contra descriptivism, Kripke and Putnam purportedly demonstrated that the relation between primary and secondary properties is noncontingent, via the intuitive plausibility of the following principle:

(ES) Essentialism about kinds: for any (general) natural kind T , T is individuated by the same set of essential properties $\{P_1, \dots, P_n\}$ across all possible worlds.⁴⁰

³⁸ For instance, see the main expository works on essentialism especially Kripke (1980), Putnam (1983) and Burge (1979).

³⁹ According to this widely held model of natural kinds, the properties characteristic of any kind K come in two grades, one more fundamental to the other. The more fundamental grade are called primary nomic properties, the less fundamental secondary nomic properties. The point of the model is that the primary nomic properties 'underlie' or 'give rise' to the secondary properties; e.g. we can explain the nomic regularities displayed by the superficial properties of gold by invoking its atomic structure. Thus, this model also implies a particular ontological relationship between property levels: primary properties are fundamental existents and secondary properties simply supervene upon them. See Colin McGinn, *A Note on the Essences of Natural Kinds*, *Analysis*, 35.6, June 1975, pp. 178-180.

⁴⁰ For a similar formulation, see Andrew Bird, 'Necessarily, Salt Dissolves In Water', *Analysis* 61 (4) 2001, pp. 267-274

Where the essential properties individuating *T* are understood as primary nomic properties which “realize” or are “responsible for” secondary nomic properties. As a result, secondary nomic properties are ontologically derivative from their primary base properties, and thus are simply *identified* with them. From this understanding of essentialism, then, it is a short step to rigidity: true property identities are (a posteriori) necessary, and thus if secondary properties are identical to primary ones, they are necessarily coextensive (rigid).⁴¹

Turning to Kornblith, it is easy to see here why such a property essentialism about natural kinds provides the foundation for a type of semantic reductionism: if the referential content of ‘knowledge’ is necessarily coextensive (rigid) with the relevant set of physical kind terms – i.e. coextensive across all possible worlds – then one has reason to think that intension and extension collapses or deflates when it comes to ‘knowledge’ (in the way that intension and extension collapse for proper names on the non-descriptivist, Kripkean view) and that it functions as a semantic primitive.⁴² Thus, reductionism about ‘knowledge’ can be defined as such:

⁴¹ More specifically, for Water to have a set of secondary properties (clear, potable....) is *just for it* to have a particular set of primary properties (say the micro structural composition H₂O). Thus, Water and H₂O are identical (necessarily coextensive) thus and related rigidly: there is no possible world in which Water isn’t H₂O and conversely, there is no possible world in which H₂O lacks the secondary properties characteristic of water. See McGinn, pg. 181. Also, consider the following passage from Kripke’s *Naming and Necessity*, page 320-321:

“In particular, then, present scientific theory is such that is its part of the nature of gold as we have it to be an element with the atomic number 79. (We may also in the same way, then, investigate further how colour and metallic properties follow from what we have found the substance gold to be: to the extent that such properties follow from the atomic structure of gold, they are necessary properties of it, even though they ... are not part of the meaning of ‘gold’ and were not known with a priori certainty).”

⁴² Intension and extension collapse in the case of necessary coextensions because in such cases we *cannot* prize apart descriptive and referential content. That is, if ‘knowledge’ = ‘physical property *p*’, then there is no possible world *W* where the descriptive content of ‘knowledge’ is associated with an alternative extension precisely because this descriptive content *just is* (ontologically) its referent. In other words, the constituent of the proposition “knowledge’ = ‘physical property *p*” is simply the physical object which necessarily “gives rise” (for reasons explained in a footnote above) to the secondary nomic properties which are associated with ‘knowledge’. Thus, it is clear that the deflationary, anti-descriptivist semantics

(RK): The natural kind-term 'knowledge' is a (directly denoting) semantic primitive which is necessarily coextensive with the same set of essential physical properties $\{P_1, \dots, P_n\}$ across all possible worlds.

Summing up this account, the kind knowledge simply *is* a set of physical base properties; a set which we then identify (via ostension) with the referent of the natural kind-term 'Knowledge' through empirical investigation.

2.4.2 *The Multiple Realizability Objection*

Understanding how property essentialism is a form of reductionism opens the way to my first criticism of Kornblith's epistemology. What I will contend here is the following: given that the natural kinds account of knowledge is reductionistic, it is vulnerable to an objection based upon the multiple realizeability (at the physical level) of epistemic properties.

First, let us review. If I am correct in explicating his position, Kornblith is committed to a metaphysical essentialism such that knowledge, as a natural kind, is *identical* to the physical kind which realizes it. Turning to the semantic side, Kornblith uses this putative kind identity to underpin a reductionism about epistemic kind-terms such that higher-order (i.e. functional) epistemic predicates, construed in an intentional, mentalistic vocabulary⁴³, are necessarily coextensional with lower-order physical predicates. And it is this semantic-metaphysical move, I emphasize, which provides the

underpinning Kornblith's reductionism about knowledge relies upon prior ontological property identifications. For more, see McGinn (1975), Boyd (1988), Kripke (1980).

⁴³ By "mentalistic", I mean that epistemic predicates are couched in the vocabulary of what is known as "folk" (belief/desire) psychology. That (unreduced) epistemological vocabulary is mentalistic is a relatively uncontroversial point, and I shall not belabour it presently. For more, however, see Tye (1991), Block (1997),

foundation for Kornblith's methodological sanction on the a priori.⁴⁴ As a result, if it could be determined that the kind identity in question *is false*, then justification for skepticism about the a priori would be undermined, the semantics of knowledge as an a posteriori natural kind *requiring* this property identity.⁴⁵

I argue that the multiple realizeability objection gives us excellent reason to believe that this kind identity is in fact false. The objection runs as follows. For reductionism to be true, the kind-term 'knowledge' must be necessarily coextensional with a particular physical kind-term '*P*'; which in turn requires a true property identity between the kind denoted by 'knowledge' and its physical realizer *P*. But the kind denoted by 'knowledge' is compatible with *multiple base properties* (*P*, *Q*, *R*, ..., *n*); implying that it *can't be* identical with *P*. So 'knowledge' *isn't* necessarily coextensional with '*P*', and thus Kornblith's reductionism about knowledge is false; it is not a structured natural kind with an individuating essence like 'gold' or 'water'. Construed in abbreviated form:

- (1) [The multiple realizeability thesis] Epistemic property-kinds such as knowledge are multiply realizable.
- (2) [The anti-identity thesis] If epistemic kinds such as knowledge are multiply realizable, then they are not identical to only one set of physical kinds.

⁴⁴ As I made clear in the previous section; Kornblith's methodological sanction on a priori methodology turns on his argument to the effect that knowledge (and other epistemological kinds) are discovered a posteriori; implying that none of our putative a priori beliefs about the phenomenon are justified. And knowledge as an a posteriori phenomenon, for reasons discussed elsewhere, requires property essentialism and the associated semantic machinery. Thus, the overall structure of Kornblith's negative argument stands or falls on soundness of his natural kinds argument for property essentialism about knowledge.

⁴⁵ As per. my discussion on the previous section, ontologically significant reduction requires the reduction of higher-level properties and this in turn requires that they be *identified with* (complexes of) lower-level properties. Identity, of course, requires (at minimum) an appropriately modalized coextensivity. See Jaegwon Kim, *Multiple Realization and the Metaphysics of Reduction*, Philosophy and Phenomenological Research, LII (1), March, 1992.

(3) Therefore, if epistemic kinds such as knowledge are not identical to physical kinds, then epistemic kind-terms are *not* necessarily coextensional with physical kind-terms.⁴⁶

To expand upon this rather general formulation, I want to look at two different variations of the argument, one providing empirical motivation for the conclusion drawn, the other providing conceptual motivation. The empirical variant is derived from an argument made by Hilary Putnam⁴⁷, and the conceptual variant is derived from an argument made by Jerry Fodor⁴⁸.

Putnam's Argument

Putnam objects to reductionism about mental properties empirically, by citing evidence regarding neurophysiological variety in higher organisms. For the fact that different species of higher organisms display highly variegated physical neurological structures, Putnam argues, coupled with the strong likelihood that they instantiate functionally indistinguishable mental properties (such as pain) implies that reductionism is empirically false.

Described in Putnam's own words:

Consider what the brain-state theorist [reductionist] has to do to make good his claims. He has to specify a physical-chemical state such that *any* organism (not just mammal) is in pain if and only if (a) it possesses a brain of suitable physical-chemical structure; and (b) its brain is in that physical-chemical state. This means that the physical-chemical state in question must be a possible state of a mammalian brain, a reptilian brain, a mollusc's brain (octopuses are mollusc, and certainly feel pain), etc.⁴⁹

⁴⁶ For a similar argument, see John Bickle, "Multiple Realizability", *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/multiple-realizability/>. Also see Terence Horgan, "Nonreductive Materialism and the Explanatory Autonomy of Psychology", in Steven J. Wagner & Richard Warner (eds.), *Naturalism: A Critical Appraisal*. University of Notre Dame Press (1993).

⁴⁷ Hilary Putnam, 'Psychological Predicates', in W. H. Capitan and D. D. Merrill, (eds.), *Art, Mind and Religion*, University of Pittsburgh Press, 1967.

⁴⁸ Jerry Fodor, 'Special Sciences, or the Disunity of Science as a Working Hypothesis', *Synthese*, 28, 1974, pp. 97-115.

⁴⁹ Putnam, 1967, 436.

If every psychological property is, according to reductionism, necessarily identical to some underlying physical-chemical property, then the antireductionist only needs to find one psychological property that is applicable to all of these creatures but whose physiological realizing base properties differ. And according to Putnam, (given neurobiological structural diversity) it is “overwhelmingly likely”⁵⁰ that such a psychological property can be found, thus implying (by induction) the falsity of reductionism.

Importantly, I think that an analogous version of Putnam’s empirically motivated multiple realizeability argument is quite effective against Kornblith’s reductionism about knowledge. This is for the following reason. Given the very plausible assumption that epistemic states just are species of psychological states, it seems overwhelmingly likely that the wide variety of organisms which realize epistemic informational states of the kind Kornblith countenances in fact do so via diverse neurophysiological base properties. For example, consider the higher-order (i.e. functional) state associated with the epistemic kind ‘justified’. It seems empirically likely that a wide variety of (actual and possible) organisms, endowed with very different physical cognitive architecture to our own, could instantiate the functional state associated with this epistemic kind. And if this is the case, it follows that epistemic kinds such as knowledge and justification are multiply realizable, with the further implication that Kornblith’s claim of a theoretical identity between epistemic and physical kinds (and hence the claim of a necessary coextension between kind-terms ‘knowledge’ and ‘*p*’) is *false*.

Fodor’s Argument

⁵⁰ Putnam, 1967, 437. See also John Bickle, *Psychoneural Reduction*, MIT Press, 1998, pp. 116-117.

The second variant of the multiple realizeability argument was originally articulated by Fodor. He claims that we have reasons, independent of empirical findings, to reject reductionism.

The argument begins with the following observation: if reductionism is true, then it follows that the lawlike regularities which can be captured by higher level psychological kind-terms (the focus of the “special sciences”) can be redescribed or mapped one-one *salva veritate* (given proper bridge laws) into lawlike regularities of a lower level reducing set of physical kind-terms. That is, according to reductionism, coextensive kind identity implies that very event which falls under a law of psychology *also* falls under a law of physics⁵¹:

If reductionism is true, then every kind is, or is coextensive with, a physical kind... this follows immediately from the reductionist premise that every predicate which appears as the antecedent or consequent of a law of a special science must appear as one of the reduced predicates in some bridge law, together with the assumption that the kind predicates are the ones whose terms are the bound variables in proper laws.⁵²

The problem, argues Fodor, is that the possibility of multiple realization⁵³ *precludes* the subsumption of higher level psychological laws to lower level physical laws. Why? Because if a particular psychological kind can be realized by multiple base kinds, then one-one mappings of kind-terms (theoretical coextension) will not be possible. At best, “wildly disjunctive” laws will obtain; laws such as:

(L1): $(K \leftrightarrow (P1 \vee P2 \vee P3 \vee P4 \vee Pn \dots))$

⁵¹ This is because higher-level kind-terms, given necessary coextension, express a biconditional relation with lower-level physical kind-terms such that $K \leftrightarrow P$. Thus, they can be intersubstituted while preserving truth-value. This is how “classical” reductionist theories were often construed. For more, see Nagel (1961).

⁵² Fodor, 1974, pp. 132-133.

⁵³ Here Fodor cites neuroplasticity data taken from the work of Karl Lashley in his (1974) paper. Also see his *‘Language of Thought’*, Bradford Books/MIT Press, 1975.

However, (and this is the key point), the disjunctive consequent of L1 *is not* a kind predicate (i.e. does not range over a distinct entity), and the entire expression is *not* a law of any current physical science; implying that reductionism about psychological kinds is untenable.

To understand the full upshot of this argument, it is necessary to understand why Fodor prohibits the reduction of psychological laws to an open-ended disjunctive set of lower-level physical laws. The primary reason, it appears, has to do with nomic projectability.⁵⁴ The idea here, I think, is that we can't reduce a higher level psychological kind to a set of disjunctive physical properties because such disjunctive sets *fail* to pass the projectability test, suggesting that they *fail to range over an actual lawlike kind through which explanatorily adequate generalizations about the reducing domain could be secured* (i.e. they are "gerrymandered" kinds).⁵⁵ Thus, contends Fodor, psychological kinds, as extensionally equivalent to an indefinite variety of base properties (i.e. they are "cross-classifying"), maintain an explanatory autonomy vis-a-vis

⁵⁴ That is, projectability, according to Kim, (1992), pg 11: "...is the ability to be confirmed by observation of positive instances such that any generalized conditional of the form 'all *F*'s are *G*' can be confirmed by the *exhaustion* of the class of *F*'s – that is, by eliminating all potential falsifiers... Lawlike generalizations, however, are thought to have the following further property: observation of positive instances, *F*'s that are *G*'s, can strengthen our credence in the next *F* being a *G*. It is this kind of instance-to-instance accretion of confirmation that is supposed to be the *hallmark* of lawlikeness; it is what explains the possibility of confirming a generalization about an indefinitely large class of items on the basis of a finite number of favorable observations."

⁵⁵ Here is how Kim (1992), pg. 10, puts the point: "...if *M* is identified with non-kind *Q* (or *M* is reduced via a biconditional bridge principle " $M \leftrightarrow Q$ ", where *Q* is a non-kind), *M* could no longer figure in special science laws; e.g. the law " $M \rightarrow R$ " would in effect reduce to " $Q \rightarrow R$ ", and therefore loses its status as a law on account of containing *Q*, a non-kind. So, in the epistemic case, reduction of the property 'knowledge' to a disjunction of physical properties would imply that we would no longer be able to accurately generalize over the property, because it would not support robust counterfactual explanation. For more, see Goodman (1963).

their lower level physiological correlates; an autonomy which establishes the falsity of reductionism.⁵⁶

As in the previous section, I believe that we can run a parallel variant of the argument against Kornblith's reductionism about knowledge. The reason is simple: given the plausible assumption that propositional attitudes about epistemic matters just are a type or subclass of higher-level psychological (intentional/doxastic) kinds⁵⁷, we have as much reason (at least short of an argument establishing the relevant disanalogy) to believe that the former are going to be (in the words of Fodor) as "wildly disjunctive" as the latter. That is, it seems justified to think that if epistemic states are, (like other psychological states), compatible with an open-ended number of base realizers, then they will fail, on conceptual grounds, to correspond to or range over an actual natural kind through which explanatorily adequate (projectable) reductive generalizations about the epistemic domain can be secured. For instance, let us consider the epistemic propositional attitude "justified belief". If Fodor is correct and psychological kinds correlate with a disjunctive set of base realizers, then it is quite clear that any attempt to reduce or explain away "justified belief" in such a manner is going to eliminate the projectability of the original epistemic generalization. Why? Because in such a case, instead of having a functionally specified state which would support general inferential predictions, we would instead be left with a set of disparate physical kinds over which no projections could be made (due to physical dissimilarity). Like other theoretical kinds

⁵⁶ Thus Fodor, 1974, pg. 113: "There are special sciences [viz. psychology] ... because of the way the world is put together: not all kinds (not all the classes of things and events about which there are important, counterfactual supporting generalizations to make) are, or correspond to, physical kinds." We could, perhaps, simply go eliminativist about psychological properties in the face of this argument. But this is an implausible option with little relevance to my current topic; so I will not dwell on it.

⁵⁷ I take it that this isn't a controversial point: what, other than token intentional states, could epistemic states be?

embedded in proportional attitudes, epistemic kinds are functionally realized, and thus resist reduction.

2.4.3 The Argument from Psychological Explanation

The second anti-reductionist argument I would like to make is rooted in some remarks made by the psychologist Zenon Pylyshyn.⁵⁸ Pylyshyn defends the indispensability of a higher order cognitive (or intentional) vocabulary to the discipline of psychology by arguing that a reductionistic account of human action is unable to capture theoretically indispensable generalizations.

It is generally agreed upon that one of the primary tasks of science is to capture lawlike generalizations in nature as a means of supporting inductive inference over a particular domain; where the generalizations in question are relative to a particular set of entities and a particular theoretical vocabulary. What Pylyshyn maintains is that even if psychological terms were, in all cases, extensionally equivalent to a set of physical terms, we could *not* reduce or deflate the vocabulary of psychological laws to the vocabulary of physical/biological laws. Since the upshot of his objection turns on an example, let me summarize it.

Folk Psychology and Explanatory Power

Take a commonplace occurrence such as a pedestrian witnessing a car accident. Seeing that people are hurt, the pedestrian runs to the phone booth and dials a '9', followed by a '1'. What will she do next? And why did she run to the phone booth? As John Bickle articulates:

⁵⁸ Zenon Pylyshyn, *'Computation and Cognition'*, Cambridge, MA: MIT Press, 1984.

When one recognizes that an emergency has occurred, one typically wants to communicate that information to those who can help the victims. One way to communicate such information efficiently is to use the emergency telephone number. Most adults in our culture know that the emergency telephone number is 911. So when a person knows that he has successfully dialed 9 and 1 in this situation, [it is clear that] he will intend to dial another 1. When it comes to telephone dialing, people tend to dial the number they intend to dial. The connection between *the perceived situation* and *the action* is systematic. There is a generalization between perceived situation and action here for the capturing, and this generalization extends beyond merely this situation to the equivalence class of all situations involving collections of beliefs and desires whose contents are similarly related.⁵⁹

It is no coincidence that we can accurately predict that the person in question will intentionally dial another '1' in this scenario. For we recognize this sequence as a particular instance of a general folk psychological generalization which counterfactually obtains between a desired state of affairs *D* (viz. 'to help', 'to phone 911', etc.) and a consequent action-type *A* intended to bring about *D* (viz. 'going to the phone booth', 'picking up the phone', 'dialling a '1'); where this psychological generalization ranges over all equivalent situations in which an agent desires a particular state of affairs, and believes that intending a particular action-type will bring about this state of affairs. That is:

(G): $(\forall S) (\forall D) (\forall A) (\text{If } S \text{ wants } D \text{ and believes that doing } A \text{ will bring about } D, \text{ then } ceteris\ paribus\ S \text{ will do } A).$ ⁶⁰

Both the scenario in question and generalization (G) are obvious to us because we grasp the psychology of belief-desire intentional states: we know that when one desires a particular end, and knows what action-type will bring it about, one will generally perform this action-type.

Now to the actual objection. Pylyshyn argues that we *cannot* capture important folk psychological generalizations such as (G) in a lower level neurophysiological or

⁵⁹ John Bickle, *Psychoneural Reduction*, pg. 152.

⁶⁰ See Terence Horgan, "Nonreductive Materialism", pg. 298.

biological vocabulary. That is, when the predicates of a particular psychological generalization are reduced to the predicates of a particular neurophysiological or biological vocabulary, *the generalizations captured under higher level psychological description disappear*. This is because of the “projectability problem”: under the assumption that any particular psychological predicate is correlated with an open-ended number of physical realizers (i.e. the predicates “cross-classify”), closed (non-disjunctive) *ceteris paribus* psychological generalizations (in this case, what is expressed by (G) above) *cease to be projectable* when reduced to the neurophysiological level. Thus, and this is Pylyshyn’s key thesis, important generalizations couched in the former vocabulary such as (G) are simply missed or eliminated when the descriptive predicates of intentional psychology are recast in the vocabulary of neurophysiology:

The neurophysiological story misses the most important psychological generalization involved: regardless of how a person learns the emergency number, regardless of how she comes to perceive the situation as an emergency, and regardless of how the person’s limbs are moved in dialing, a single, general principle is implicit in the... sequence.⁶¹

Hence, *only if* we retain the higher-level intentional vocabulary of folk psychology can physically unrelated processes be organized into a (belief-desire-action) equivalence class over which robust (projectable) counterfactual generalizations be fashioned, for two levels of explanation can extensionally converge without providing equal predictive capacities (because of differences in intensional content).⁶²

⁶¹ Pylyshyn, 1984, pg. 11.

⁶² Note that the argument from Psychological Explanation, although invoking the premise of multiple base realizers, is dialectically distinct from the multiple realizeability objection for the following reason. Even if it was unproblematically possible, metaphysically speaking, to identify psychological predicates with a set of disjunctive realizers, this would not, at least on the face of it, settle the *epistemological* question of explanatory efficacy or relevance. That is, Pylyshyn’s argument for the indispensability of higher level folk psychological explanation seems to hold weight even if at some level an ontological reduction was possible.

So, the argument concludes, a reductionism about psychological predicates violates one of the primary tenants of scientific inquiry: capture *all* capturable generalizations. When folk psychology is deflated, significant counterfactual regularities are *missed*. And as a result, we have very good inductive grounds for rejecting reductionism (as a means of maximizing the predictive capacities of theory).

Applying the Argument to Kornblith

I believe that Pylyshyn's objection to reductionism can be wielded against Kornblith's particular brand of epistemological reductionism. More specifically, I will contend below that one upshot of the argument from psychological explanation is that puts Kornblith face to face with a constructive dilemma: either (I) maintain reductionism, the consequence being the projectability problem, or (as a means of avoiding the projectability problem) reject reductionism (II) therein undermining motivation for the methodological sanction on the a priori. This dilemma, I will argue, renders Kornblith's natural kinds account of knowledge highly implausible.

I. The First Horn

First, note that many generalizations captured by the vocabulary of intentional psychology express propositional attitudes that are explicitly *epistemological*. That is, doxastic intentional states like '*believes*', '*knows*' and '*has evidence for*' express an epistemic relation between a propositional attitude and a putative state of affairs; an intentional relationship which is sufficiently robust to support *ceteris paribus*

psychological generalizations. Here is an example of such a generalization involving the propositional attitude ‘has evidence for’.⁶³

(EG): $(\forall S) (\forall E) (\forall B)$ (If a subject S has evidence E for her belief B , then S will be disposed, *ceterus paribus*, to affirm B when prompted).

Second, recall the basic import of Kornblith’s reductionism: for any epistemological predicate P expressed in a (higher level) psychological vocabulary, P is necessarily coextensional with a (lower level) predicate Q expressed in a physical vocabulary.⁶⁴ So, when we turn to a higher level *ceterus paribus* generalization like (EG), it appears that Kornblith would be committed to a position such that necessarily, for each higher level psychological term of the generalization (EG), there is a lower level physical term such that the latter is necessarily coextensive with the former.⁶⁵

If I am correct here, then it is easy to see how Pylyshyn’s objection can be brought to bear on Kornblith. For if epistemological predicates expressed in a psychological vocabulary reduce to lower order physical predicates (which follows from essentialism), and we assume the very plausible premise of open-ended base realizers, then the ‘projectability problem’ arises for Kornblith. More specifically, if the terms of a generalization like (EG) are correlated with multiple physical realizers, then (as Pylyshyn demonstrates with an unspecified psychophysical reductionism) the generalization in question simply *disappears* when reduced to the physical level; at least insofar as its conjunctive physical correlates don’t support projectable counterfactuals. This is a significant problem for Kornblith’s natural kinds account. For if the reductive view of

⁶³ For more on epistemic states as propositional attitudes, see Horgan (1993).

⁶⁴ Where ‘coextension’ is understood in ontological terms as per. section 2.3.1.1; i.e. P is necessarily (*de re*) coextensional with Q according to Kornblith because Q denotes a set of primary nomic properties which are “constitutive of” and “give rise” to the secondary nomic properties denoted by P .

⁶⁵ See footnote 42 above and Kim, (1992), Horgan, (1993).

epistemic properties cannot capture higher level generalizations such as (EG), then it lacks explanatory power vis-a-vis its initial target *explanandum* and thus appears to be self-defeating.

II. The Second Horn

Can Kornblith's account of knowledge successfully avoid the projectability problem if modified? I argue that it cannot do so without usurping motivation for the prohibition on a priori analysis. This becomes clear when the following two points are emphasized.

First, as I made clear above, recall that the essentialist strategy for prohibiting the a priori is at the root of the projectability problem. This point becomes apparent when it is understood that reduction of higher level epistemic terms to lower level physical terms requires that the kinds denoted by each sort of term are identical (i.e. a relation of necessary coextensivity).⁶⁶

Second, recall that adopting property essentialism *also* underlies premise (D) of the Natural Kinds subargument:

(D) The referential content of natural kind terms is discovered empirically by investigating their locally-determined essence [assumption via causal semantics + kind essentialism].

The stipulation of essentialism in (D) makes it the crucial premise in the subargument for the following reason. In identifying epistemological kinds with unique physical kinds, premise (D) *provides the metaphysical underpinning* for the semantics of the causal theory of reference.⁶⁷ And it is adopting a causal theory of reference which justifies

⁶⁶ See footnotes 21-23 above.

⁶⁷ More specifically, it is generally acknowledged that the semantics of the causal theory of reference, in order to deflate descriptive to referential content (in the "Millian" sense), relies upon an assumed (metaphysical) essentialist premise to the effect that the secondary properties *S* of some natural kind *K* are

Kornblith's argument to the effect that the business of understanding the semantic content of the term 'knowledge' is an entirely a posteriori affair.

Crucially, then, when these two points are combined, we can see that if Kornblith were to try and circumvent the projectability problem by modifying premise (D) (by adopting an alternate theory of reference), he would be *unable* to maintain his proscription against a priori analysis. Why? Because if I am correct, justification for Kornblith's Exclusionary Argument would *fall away* without justificatory support of the essentialist thesis (and the causal theory of reference as its semantic analogue). More specifically, only through the essentialist thesis in question can Kornblith repudiate a functional (higher order) view of epistemic properties; and such repudiation is necessary for the Exclusionary Argument to have any traction whatsoever.⁶⁸ Thus, modifying the relevant premise of the Natural Kinds subargument in the face of the projectability problem doesn't seem like a live option for Kornblith, lest he was to abandon his overall metaepistemological aims.

The Constructive Dilemma

If my arguments presented in (I) and (II) above are sound, then Kornblith is placed in an awkward dialectical position. While the projectability problem poses severe problems for a reduction of epistemic properties to physical correlates, it appears that the reductive component (i.e. premise (D)) of Kornblith's Natural Kinds subargument can't be

identical to a set of primary properties *P* which constitute or "give rise to" *S*. For more details, see footnote 23. above and Putnam (1975).

⁶⁸ This is, roughly, for the following reason. Because functional properties are identified by their descriptive content, they are thus amenable to a priori conceptual analysis (that is, we can know, a priori, whether a possible world contains a particular functional kind by simply examining the meaning of the relevant terms and intensions).

modified without undermining justification for the key premise of his Exclusionary Argument.

Thus, we can run the following constructive dilemma against the natural kinds account of knowledge:

- (1) If Kornblith retains essentialism about natural kinds, then his account will fail to capture important epistemological generalizations [from (I)].
- (2) If Kornblith rejects essentialism about natural kinds, then his account will fail to provide justification for his proscription on a priori analysis [from (II)].
- (3) Either Kornblith retains essentialism or drops essentialism.
- (4) Therefore, either his account will fail to capture important epistemological generalizations, or it will fail to provide warrant for a sanction on a priori analysis.

This line of reasoning, I think, gives us a good case for rejecting Kornblith's subargument for essentialism, and thus rejecting the exclusionary argument against a priori justification which it supports. For my constructive dilemma provides good reason for rejecting (D) of the subargument, and hence *leaves us with no good reason to accept* premise (5) of the exclusionary argument. In particular, it leaves us with no good reason to accept premise (5)'s presupposition that 'knowledge' (or other epistemic kind terms such as 'justification', etc.) denotes a natural kind. (More specifically, Premise (5) says: "*A priori* analysis cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain." This is equivalent to: "(a) The referential content of the term contained within *the* target domain of a priori analysis is a natural kind, and (b) *a priori* analysis cannot illuminate this referential content." I reject (a) here, but not (b). And the point is that, given the soundness of my constructive

dilemma, we can now see that Kornblith has no good reason to prevent us from rejecting (5) overall by rejecting the (a) bit of it).⁶⁹

So, whether Kornblith embraces one horn or the other, his Natural Kinds subargument appears to be undermined. And with this subargument undermined, the core Exclusionary Argument loses its force.

2.5 The “Starting Points” Objection⁷⁰

My final criticism of Kornblith presents a further problem for the natural kinds account of knowledge. More specifically, I will argue the following. Because the kind knowledge (unlike paradigmatic natural kinds like gold or water) lacks observable properties, Kornblith is forced to rely upon a priori application intuitions in order to identify cases of it. However, because his Empirical Method commits him to holding such application intuitions unreliable, we have no reason to believe that he has successfully illuminated his target explanandum. Thus, assuming for the sake of argument that the proper targets of inquiry are natural kinds, Kornblith’s account of knowledge is open to the very same objections as his “armchair” rivals; implying a form of epistemological scepticism.

2.5.1 Kornblith, Conceptual Analysis and the Empirical Strategy

In order to set up the objection, we need to briefly review both Kornblith’s rejection of a priori conceptual analysis vis-a-vis epistemological kinds and the empirical method he puts in its place. The root of Kornblith’s rejection follows from his view of

⁶⁹ Thanks to David Matheson for suggestions here.

⁷⁰ The title of this objection was chosen in homage to George Bealer (1996), who makes a similar argument in opposition to Quinean empiricism.

epistemological inquiry as directed at extramental natural kinds. More specifically (given that he is correct about Extra-Mentalism) he *denies that conceptual analysis can be justified on the basis of application intuition.*⁷¹

Given the nature of the target phenomenon, then, the way in which epistemological methodology must proceed, according to Kornblith, resembles the methodology used to study other extramental natural kinds (like geological kinds):

When we appeal to our intuitions about knowledge, we make salient certain instances of the phenomenon that need to be accounted for, and that these are genuine instances of knowledge is simply obvious, at least if our examples are well chosen. What we are doing, as I see it, is much like the rock collector who gathers samples of some interesting kind of stone for the purpose of figuring out what it is that the samples have in common. We begin, often enough, with obvious cases, even if we do not yet understand what it is that provides the theoretical unity to the kind we wish to examine.⁷²

How are we to single out natural kinds, if not by conceptual analysis? Here Kornblith appeals to the causal theory of reference: we “point to” a salient empirical sample of the relevant phenomenon, and then ostensively “baptize” it with a general kind term. We then proceed to scientifically investigate the identified kind, in order to discover its essential nomic properties. That is, we use the following two stage strategy for picking out natural kinds:

The Empirical Strategy (ES):

(1) *The Initial Designation*: A commonplace sample is ostensively baptized with a natural kind term *K*

(2) *Empirical Discovery*: Investigation reveals what essential physical properties *O* unites all samples of kind *K*.⁷³

⁷¹ Kornblith, 2002, pp. 10-11.

⁷² Ibid, pg. 10-11.

⁷³ This isn't a commitment rendered explicit by Kornblith, but it is implicit within his discussion. See his 2002, chapters 1-2.

Importantly, Kornblith argues that our a priori application intuitions do not provide any real bearing on (ES) because they have no way of illuminating the deep physical structures unifying natural kinds. Only when we begin to empirically examine “samples” do we gain insight into the unifying properties underlying the kindhood of the phenomena in question.

2.5.2 The Disanalogy Premise and Knowledge

Kornblith claims that we can ostensibly “pick out” samples of our target natural kind and then proceed to investigate it empirically. I argue here that (i) it is doubtful that knowledge can be picked out by ostension in a way analogous to paradigmatic natural kinds like water or gold. Furthermore, (ii), I argue that this disanalogy has significant repercussions for his actual method of identifying putative cases of knowledge *contra* (ES).

i. The Disanalogy Premise

Let us return to the “Initial Designation” stage of Kornblith’s (ES). While it seems well suited for discovering samples of rocks, water or gold, it seems less than well suited to the investigation of knowledge. For how could we find and designate a “commonplace sample” of knowledge? Where in nature would we look? Is knowledge a collection of cells? A bundle of C-fibres? A genus of organism?⁷⁴ Thus, on further examination Kornblith’s analogy to the rock collector highlighted above breaks down into a *disanalogy* – for it doesn’t seem as if we could ostensibly “point to” knowledge:

⁷⁴ Note that this problem of identification is orthogonal to what Devitt calls the “Qua Problem” for causal theories of reference. See Devitt (1980).

(DP) *The Disanalogy Premise*: In contrast to water, or gold, knowledge does not have observable properties which we could ostensively “point to” or “pick out”.

The reason for the disanalogy, furthermore, seems clear: unlike water or gold, knowledge *lacks* a set of observable macro properties (in other words: an empirical “mode of presentation”) through which we could associate a kind term with an extensional reference class via an initial baptismal act.⁷⁵ In other words, we can’t just look into the world, point and say “That *such-and-such* *x* shall be designated with the kind term ‘Knowledge’”.

ii. *Fixing the Reference of Knowledge*

Assuming that the Disanalogy Premise is correct, Kornblith can’t simply “point to” and ostensively designate commonplace samples of knowledge. However, it is important to note that he does, in fact, succeed in individuating a reference class for the relevant term. For consider here a basic description of Kornblith’s reliabilist account of knowledge:

(K) *Knowledge*: A reliably produced informational state or belief (relative to a particular external environment).⁷⁶

The relevant question thus becomes: how *could* Kornblith justify an identification of knowledge with reliably produced belief, if he can’t do so, given the relevant disanalogy outlined above, by causal/ostensive means?

⁷⁵For context, think of how the Empirical Strategy would work in the case of a paradigmatic natural kind like gold. (1) We initially point to and designate a descriptive set of macro properties, i.e. “shiny, gold-colored, malleable...” with the general term ‘Gold’. (2) The kind is then investigated in order to find out its underlying essence (i.e. the atomic number 79). Why does this work in the case of ‘Gold’? Because as a chemical kind it has a set of observable properties (supervening upon its primary nomic properties) from which the reference class can be individuated in an a posteriori manner. Knowledge, in contrast, doesn’t seem to have this descriptive set of macro properties which could serve to individuate the relevant reference class via a posteriori investigation.

⁷⁶ See Kornblith, 2002, pp. 63-69.

I argue that Kornblith's justification for this identification implicitly proceeds by means of a priori analysis with its reliance on application intuitions, and indeed, I believe, *must* proceed in this manner. The reason (roughly stated) is this: because the complex property denoted by the description 'reliably produced belief' is functional rather than physical,⁷⁷ the only possible way Kornblith could secure the property identity in question would be through an a priori assessment of the intensional content of the concept knowledge (vis-à-vis a set of hypothetical scenarios).⁷⁸ Thus, despite his explicit endorsement of (ES), Kornblith can't in practise actually rely upon a causal, a posteriori account of epistemic properties in order to justify his theory of knowledge; this being impossible given the particular manner in which the reference of the term 'knowledge' must be secured.

2.5.3 Extramental Kinds, Starting Points and Relevancy

If Kornblith does in fact implicitly rely upon a priori analysis in order to justify his theory of knowledge, this is a very significant finding. For in what follows I argue that such a methodological reliance is at odds with his metaphysical commitments to extramental natural kinds.

Recall that Kornblith considers the target of epistemological inquiry to be an extramental natural kind, a kind which displays projectable homeostatic regularities. So

⁷⁷ Knowledge is functional because, roughly, it denotes via or through a descriptive set of causal powers which do not reduce to a set of structured (i.e. microphysical) properties.

⁷⁸ The process for arriving at such an account of 'Knowledge' would go as follows. First, we would analyze the intuitive application (extension) of our concept of knowledge in a hypothetical scenario containing reliable informational states. This would provide the evidential input for an a priori function which would generate as output a *prima facie* justified belief about the intensional content of our implicit concept of knowledge. Only if we intuited that all and only those scenarios where reliable informational states are instantiated were scenarios where knowledge is instantiated would we have a priori evidence for a justified belief regarding the coextension of the relevant predicates.

far, so good. But now the problem: if Kornblith holds that knowledge has an extramental essence, yet (implicitly) relies upon a priori analysis in order to initially determine the content of the concept 'Knowledge', then we have *no reason to think that his account of the phenomena will closely correspond to its extramental essence*. Why? Because if knowledge is really an extramental phenomenon (which is, I note, a thesis which Kornblith explicitly endorses) then we don't have any reason to think that somehow, a priori, he is zeroing in on or referring to the correct extramental properties. In short: our concepts can *misrepresent* the world. This point can be further cashed out by keeping in mind the disanalogy between knowledge and paradigmatic natural kinds like gold or heat. As Bryson and Alexander put it:

Note the contrast between gathering samples of water or heat and gathering samples of knowledge. Our intuitions tend to be far less uniform in the epistemic realm. This is probably because we identify water and heat through observable properties. They have a certain group of properties that uniformly influence our classificatory judgments.⁷⁹

Because we can identify the latter kinds via observational properties, zeroing in on the correct explanandum is both relatively straightforward and uncontroversial. But this isn't at all the case with knowledge: we must rely upon our intuitions about possible cases in order to determine the content of the relevant concept; intuitions which are both highly contested and controversial.

With this, the pieces are finally in place for my Starting Points objection, which proceeds as follows. As we have seen, Kornblith relies upon a priori analysis to justify his reliabilist account of knowledge. However (assuming that the proper targets of epistemic inquiry *are* extramental natural kinds) the Empirical Strategy commits him to holding that a priori intuitions are totally unreliable vis-à-vis the target explanandum.

⁷⁹ Anthony Bryson and David Alexander, "*The View from the Armchair: Responding to Kornblith's Alternative to Armchair Philosophy*," *Essays in Philosophy*: Vol. 13, (1), Article 10, 2012, pg. 15.

Thus, Kornblith *isn't justified* in holding that his preferred account of knowledge (as reliably produced true belief) successfully picks out or corresponds to a natural kind.

In other words (as a result of the disanalogy premise), Kornblith, like his primary rivals, is forced to rely upon a methodological starting point which is fundamentally ill-suited to the illumination of an empirical target phenomena. So (and this is the crucial point) it doesn't seem (at least on the face of it), that he – or anyone else for that matter – will be able to successfully illuminate the target phenomenon; which is, I add, tantamount to a strong form of scepticism. Summarized, the objection is this:

The Starting Points Objection

- (1) As per the Disanalogy Premise, Kornblith must rely upon a priori conceptual analysis in order to justify his account of knowledge as reliably produced belief.
- (2) However, Kornblith's (ES) commits him to the unreliability of a priori intuitions vis-a-vis the target explanandum (because concepts can misrepresent a natural kind).
- (3) Therefore, Kornblith can be no more justified in holding that his preferred account of knowledge, rather than those of his "armchair" opponents, successfully picks out an extramental natural kind (or has any relevance to an extramental domain of inquiry); implying a strong form of epistemological scepticism.

If this argument is sound, Kornblith's actual metaepistemology (as opposed to his official, empirical method) is henceforth open to the *very same* objections he makes to those who explicitly advocate an "armchair" approach. That is, why should we think that Kornblith's analysis of knowledge is at all (or any more than his rivals) relevant to the domain of inquiry at hand? And what about theoretical contamination? If we start off with faulty foundations, then how are we supposed to justify our later, theoretically mediated intuitions?

My purpose here is not to needlessly rehash Kornblith's methodological prohibition of the a priori, but rather to point out that because he must, in fact, implicitly

rely upon conceptual analysis, then the relevant methodological sanction acts as a sceptical *reductio ad absurdum* of his own metaepistemology. For if I am correct and there is no other way to successfully analyze the term 'Knowledge', then it isn't at all clear that we can justify our starting points (vis-à-vis a natural kind) whatsoever.⁸⁰

2.6 Conclusion

As a means to concluding this chapter, let me first summarize the general outline of Kornblith's empirical metaepistemology, and second, review the three arguments I have articulated in objection to it, and comment on the potential implications. My hope is that this summary will bring the dialectical import of the discussion into sharper focus in support of my overall aims.

Summary of Kornblith's Argument

In the broadest terms, Kornblith can be described as advocating a naturalistic version of the (ICA) principle; as one of his primary aims is to argue that a thoroughly naturalized metaepistemology must reject a priori methods outright (or, to put it in terms of the schema introduced in the first section, Kornblith argues in favour of an affirmation of (i) and denial of both (ii)-(iii)). The basis for this methodological exclusion of the a priori, furthermore, is based upon his particular view of knowledge as an extramental natural kind; it is an a posteriori, empirical phenomenon which therefore *isn't* amenable to conceptual analysis.

Here is the so-called 'Exclusionary Argument' summarized:

⁸⁰ More specifically if I am correct here then any empirical hypothesis about the reference of 'Knowledge' will be subject to an undercutting defeater, thus rendering any such hypothesis unjustified.

- (1). If epistemology is to be methodologically naturalized in a thoroughgoing manner, then it cannot retain scientifically suspect methods
- (2). Epistemology is to be methodologically naturalized in a thoroughgoing manner
- (3). Therefore, epistemology cannot retain scientifically suspect methods.
- (4). If an epistemological method cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain, then it is a scientifically suspect method.
- (5). *A priori* analysis cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain.
- (6). Therefore, a priori analysis is a scientifically suspect method.
- (7) Therefore, epistemology cannot retain a priori analysis.

As I made clear above, the lynchpin premise is (5), which, in presupposing that knowledge is an extramental natural kind, sets up the desired methodological sanction expressed by (6) and (7). Thus, Kornblith attempts to shore up this key premise by defending the thesis of knowledge as a natural kind, which I subsequently reconstructed as the following ‘Natural Kinds’ subargument:

- (A) If a term denotes a co-occurring set of properties which: (i) constitute a “Homeostatic Cluster”, (ii) this set is “projectable”; i.e. it persists through external change and figures centrally in causal/nomological explanations, then this term denotes a Natural Kind.
- (B) The term ‘knowledge’ denotes a co-occurring set of properties which satisfy conditions (i) and (ii).
- (C) Therefore, ‘knowledge’ denotes a natural kind.
- (D) The referential content of natural kind terms is discovered empirically by investigating their locally-determined essence.
- (E) If the referential content of natural kind terms is discovered empirically by the investigating their locally-determined essence, then the referential content of such terms *cannot* be illuminated by a priori analysis.
- (5). Therefore, a priori analysis cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain.

Kornblith attempts to defend his rejection of conceptual analysis by arguing that knowledge is a natural phenomenon, akin to gold or water. This allows him to then claim that the reference of the kind term 'Knowledge' is thus determined empirically, setting up the further claim (expressed in the main argument) that a priori analysis can't illuminate the target explanandum.

Objections & Implications

With Kornblith's position reviewed, let me now provide a recap of my three objections. I will then comment on why I think the objections in question effectively challenge Kornblith's desired methodological conclusions.

My first two objections directly attacked the Natural Kinds subargument. More specifically, after demonstrating that Kornblith's identification of epistemic properties with Boydian natural kinds is a form of reductionism (section 2.4.1), I raised the following two concerns:

(1) *Multiple Realizeability*: if reductionism is true, then it follows that each epistemological kind (e.g. knowledge) is type-identical to a particular physical kind. However, such epistemological kinds are compatible with multiple physical base realizers. So the identity relation fails to hold, and thus reductionism is false.

(2) *Psychological Explanation*: Pylyshyn puts forth the following antireductionist objection. According to reductionism, the projectable generalizations of a higher level mentalistic vocabulary should also be capturable when deflated into the vocabulary of the lower level reducing (physical) kinds. But given that multiple physical base kinds

can instantiate the relevant mentalistic kinds, mentalistic laws can only be correlated with a non-projectable disjunct of physical predicates. So, the relevant mentalistic generalizations simply disappear when reconstrued on the physical level.

I then argued that this objection can serve as the basis for a constructive dilemma aimed at premise (D) of Kornblith's Natural Kinds subargument: the constructive dilemma in question provides good reason for rejecting (D) of the subargument, and hence *leaves us with no good reason to accept* premise (5) of the exclusionary argument. In particular, it leaves us with no good reason to accept premise (5)'s implication that 'knowledge' ('justification', etc.) denotes a natural kind.

My third objection, rather than directly attacking the Natural Kinds subargument, instead highlights an important disanalogy between epistemic kinds as putative natural kinds and other natural kinds in order to pose a problem for Kornblith's methodology:

(3) *Starting Points*: If my disanalogy premise is correct, Kornblith is forced to rely upon a priori analysis in order to justify his reliabilist account of knowledge. However, his natural kinds view implies that a priori analysis lacks justificatory power vis-à-vis the relevant target domain. Thus, I argue (granting the natural kinds view) that Kornblith can have no more justification for thinking that his reliabilist analysis of knowledge is any more successful at illuminating the target explanandum than any of his armchair competitors, implying a form of general epistemological scepticism.

To conclude: I think that these objections give us good grounds for rejecting Kornblith's version of Incompatibilism, i.e. his reasons for embracing something like the (ICA) principle.

Chapter 3: BonJour

3.1 Introduction

The second Incompatibilist position I will consider is Bonjour's Non-Naturalism. To review, Non-Naturalism affirms an exclusionary position regarding naturalism and a priori analysis to the effect that a rejection of naturalism is entailed by any non-sceptical metaepistemology (where it is assumed that naturalism is incompatible with a priori analysis). Vis-à-vis my taxonomy, in other words, a Bonjour-style Non-Naturalist claims that to retain (ii)-(iii), we must reject (i).

My primary objective in this chapter will be to challenge both the viability and underlying motivation of Bonjour's rationalist Non-Naturalism. This will be achieved by attacking the elimination argument Bonjour uses to demonstrate the sceptical consequences supposedly entailed by naturalism. This argument, which I designate the 'Deficiency Argument', fails (or so I will attempt to demonstrate) because it relies upon an implausible (question-begging) premise. Thus, to continue my objection, Bonjour's argument for adopting non-naturalism by elimination (i.e. that a rejection of (i) is entailed by a consistent retention of (ii)-(iii)) totally lacks motivation, opening dialectical space for my own naturalistic account of a priori analysis.

I will begin (3.2) with an overview of Bonjour's Deficiency Argument for Non-Naturalism and a discussion of his positive theory of a priori justification (3.3-3.4) before moving into an extended discussion of the aforementioned objection (3.5).

3.2 Bonjour (I): Against Naturalism

As I attempted to make clear in the introductory chapter (1.4.2), Bonjour makes the case for his own non-naturalist metaepistemological position negatively. That is, through his

Deficiency Argument Bonjour proceeds by elimination to the claim that an adoption of a rationalist view of the a priori is the only possible way to avoid a crippling global scepticism.

Accordingly, my aim in this section is twofold. First, I will examine the Deficiency argument and make some important observations. Second, I will identify the key premise, and unpack the subargument he constructs in order to defend it.

3.2.1 The Deficiency Argument

My first task is to provide a more detailed overview of the Deficiency argument outlined in the introductory chapter. This will, I think, give us a very good point of entry into the dialectical contours of the relevant discussion.

As I present it here, the Deficiency Argument can be viewed as the compression of an extended metaepistemological survey and argument by elimination found within Bonjour's book *In Defense Of Pure Reason*. To recall my first chapter, it runs as follows:

The Deficiency Argument:

- (1). Either (radical) Naturalism is a minimally adequate meta-epistemological theory or Non-Naturalism is true.
- (2). If Naturalism is to be minimally adequate, then it mustn't imply scepticism
- (3). Naturalism implies scepticism.
- (4). Therefore Naturalism isn't minimally adequate.
- (5). Therefore Non-Naturalism is true.

Let me here make a couple of important exegetical observations. First, the reason I call it a "Deficiency" argument should be clear: Bonjour thinks that naturalism, as a

metaepistemological position, implies scepticism and thus fails to be minimally adequate. Second, it should also be clear that (3) is the key premise; in order for the argument to work, BonJour is going to have to make a clear case for the sceptical consequences entailed by his reading of naturalism. As such, I will examine the case he makes for (3) below, and reconstruct it as a subargument. Before I proceed to do this, however, I first need to clear up two outstanding ambiguities.

First, in the Deficiency Argument BonJour restricts the scope of his use of ‘Naturalism’ to a Quinean “radical empiricism”; that is to say the view that *all* epistemic justification ultimately derives from sense data. The reason why he excludes what we can call “moderate empiricism” (which limits the target domain of a priori justification to beliefs about one’s own mental states) from the discussion has to do with his contention that all such views implicitly collapse into a form of rationalism. And while this claim is in and of itself very controversial, it isn’t central to my purposes here and as such I shall not pursue it further.⁸¹

Second, note that in the background to this argument stands BonJour’s assumption of foundationalism. That is, he is assuming that both putative a priori and a posteriori sources of justification terminate in basic (non-inferential) belief states. And again, while I won’t pursue the issue here, a basic apprehension of this point is important for what follows.⁸²

⁸¹ Note that while BonJour’s direct target isn’t naturalized epistemology per se., but rather “empiricism” (broadly construed), it is quite clear that he sees the two doctrines as *mutually entailing*. For as BonJour depicts the dialectical situation, moderate and radical empiricism are both motivated primarily by a desire to account for the a priori in a naturalistically acceptable way, one which therein limits the scientifically suspect elements of rationalism. See Laurence BonJour, *In Defense Of Pure Reason*, Cambridge UK: Cambridge University Press, 1998, pp. 17-20

⁸² BonJour’s general commitment to foundationalism constitutes a reversal of opinion from his earlier epistemological work. In *The Structure of Empirical Knowledge* (1985), for example, he embraced and defended a coherentist account of a posteriori justification (even though in that work he suggested a

3.2.2 *The Generality Subargument*

With Bonjour's metaepistemological argument for Non-Naturalism reviewed, I now want to take a look at how he defends its key premise (3) through an argument designed to demonstrate that inference requires non-experiential sources of justification. Explicating this attack will also, I think, provide a natural point of entry into a discussion of his positive commitments.

BonJour defends premise (3) using a deceptively simple (and highly compressed) line of reasoning which I call the 'Generality' subargument. Following Albert Casullo's helpful discussion, the basic idea of the Generality subargument is this. Some beliefs are justified solely by experience; their content being limited to particular spatiotemporal objects. Either some beliefs whose content extends beyond direct experience are justified or scepticism is true.⁸³ However, the justification of beliefs extending beyond direct experience requires inference, and inferential justification is general and a priori (i.e. non-experiential). So, either Naturalism (in the relevant sense, according to which beliefs cannot be justified non-experientially) is false, or a general, radical form of scepticism is true.⁸⁴ BonJour himself develops the argument in the following manner:

... [Experiential] beliefs are particular rather than general in their content and are confined to situations observable at specific and fairly narrowly delineated places and times. The obvious and fundamental epistemological question then becomes whether it is possible to infer, in a way that brings with it epistemic justification, from these foundational beliefs to beliefs whose content goes beyond direct experience or observation: beliefs about the past, the future, and the unobserved of the present; beliefs that are general in their content or beliefs that have to do with the kind of things that are not directly observable.

foundationalist account might have to be given of a priori justification). But by the time *In Defense of Pure Reason* came out, he had abandoned his earlier coherentist view and had become a foundationalist across the board.

⁸³ BonJour, 1998, pg. 3-4.

⁸⁴ See Albert Casullo, *A Priori Justification*, Oxford UK: Oxford University Press, 2003, pp. 101-102.

And thus:

If the answer to this question is “no”, then the upshot is a quite deep form of scepticism [a radical or global scepticism]. But if the answer is “yes”, then such inferences must seemingly rely upon either premises or principles that are at least partially justified a priori. For if the conclusions of the inferences genuinely go beyond the content of direct experience, then it is impossible that those inferences could be entirely justified by appeal to that same experience. In this way, a priori justification may be seen as essential if extremely severe forms of scepticism are to be avoided.⁸⁵

The key point, I think, seems to be that a cognitive state can’t directly justify a belief whose content goes beyond that of the state in question; implying here that basic experiential beliefs (as directed at and limited to particular objects⁸⁶) can’t act as premises justifying conclusions abstracted from the particular spatiotemporal content of the original experience. As a way of clarifying the basic line of reasoning, let us consider it in reconstructed form:

*The Generality Subargument*⁸⁷:

- (1). Experience is limited to *particular* objects.
- (2). No cognitive state can directly justify a belief whose content goes beyond that of the state.
- (3). Principles of inference are *general*.
- (4). Therefore, experience cannot directly justify principles of inference.

As it stands, however, the argument might seem to stand in danger of proving too much, because it might seem that premise (2) precludes even the possibility of a priori justification of principles of inference — if it turns out that the non-experiential cognitive states (e.g., “rational insights” on BonJour’s reckoning) which supposedly serve as the basis for such justification are also limited to particular contents. So, in order to establish

⁸⁵ BonJour, *In Defense of Pure Reason*, pg. 4.

⁸⁶ Or, it seems, other a posteriori cognitive states such as memory which, though not directly via experience, ultimately depend on experientially basic cognitive states of the type BonJour is concerned with.

⁸⁷ I adopt this reconstruction from Casullo, 2003, pp. 103-104.

his desired conclusion and nothing more, Bonjour owes us a sound explanation of the relevant difference; i.e. of how supposedly non-experiential cognitive states (like “rational insights”) can have general contents, in contrast to experiential cognitive states, which can only have particular contents.

3.3 Bonjour (II): Rationalist Non-Naturalism

At this point, Bonjour has argued for the conclusion that experiential states alone cannot justify general belief in principles of inference, on the grounds that the content of these experiential states is limited to particular spatiotemporal objects. Despite this, as I have explained, the subargument is still incomplete, for Bonjour has yet to explain how a priori, non-experiential beliefs can have *general* inferential principles or properties as content.

Unpacking how Bonjour argues in favour of the claim that non-experientially justified beliefs can (in contrast to experiential beliefs) have general content will require, however, a look at his positive commitments to (what I have called) a Non-Naturalist and Extra-Mental conception of a priori justification. Accordingly, explicating and clarifying these aforementioned commitments will be my aim in this section.

3.3.1 Platonism, Mental Content and Epistemic Access

Given that Bonjour assumes the truth of these doctrines in presenting his account of a priori justification, I need to begin with a look at his Platonist theory of abstract entities and its associated mental and epistemological machinery. In what follows, I will briefly examine each in turn.

Platonism

BonJour is committed to an ontological realism about abstract universals (or “properties” in his preferred nomenclature). These properties, furthermore, while causally inert, are nevertheless instantiated by the mind in token acts of “rational insight”. Here is the passage in which he most explicitly discusses this commitment:

Certainly it is often assumed ... that rationalism is committed to Platonism. Moreover, intuitive presentations of rationalism ... surely seem on the surface to support such an assumption: such accounts are naturally formulated in terms of propositions and the properties, relations and perhaps other metaphysical items that are their ingredients, and it is at least plausible to construe these things as abstract entities of a Platonistic sort ... my own inclination is to think that Platonism is in fact the *only* tenable account of these matter, and thus I will proceed ... on that assumption.⁸⁸

So what we have here is a very robust version of the doctrine: propositions, properties and relations are all to be construed as abstracta which somehow exist outside of the spatiotemporal order of things.

Mental Content

A “thought”, in BonJour’s terminology, consists of the mental act of instantiating or exemplifying one of the aforementioned abstract Platonic universals.⁸⁹ And when I have such a thought (i.e. instantiate an abstract property), I therein “grasp” or “apprehend” the relevant property in an unmediated (and seemingly *sui generis*) manner. That is, I don’t instantiate a symbol or representation (which would only be meaningful in relation to some external object/entity and thus would only be another spatiotemporal particular); rather I am phenomenologically acquainted in a direct manner with the property *itself*:

...the universal instantiated by thoughts of triangular things is a more complex universal having the universal triangularity as one of its components ... it is a necessary, quasi-logical fact that a thought instantiating a complex universal involving the universal triangularity in

⁸⁸ Bonjour, 1998, pg. 158. Italics added for emphasis.

⁸⁹ Ibid, pp.180-186.

the appropriate way ... is about triangular things. In this way, the content of the thought would be non-contingently captured by the intrinsic character of the mental act...⁹⁰

Thought, in Bonjour's technical sense, involves the non-contingent (i.e. "quasi-logical" and "direct") instanting of a universal such as Triangularity. Thus, what Bonjour is advocating (against representationalism and LOT theories) is the so-called "narrow"⁹¹ view of mental content; i.e. the view that the contents of my thought about some abstract property are completely independent of any relation to the external spatiotemporal environment.⁹²

Epistemic Access

How is it that we access or causally interact with abstract entities? *Prima facie*, this is a pressing issue for any theory positing a Platonistic ontology. Why? Because a very plausible condition on any theory of justification (or knowledge) is "epistemic access"; the basic idea being that justification requires (as a necessary condition) the existence of some kind of reliable causal connection between the belief and the object, situation or entity to which the belief supposedly represents.⁹³

⁹⁰ Bonjour, 1998, pg. 184.

⁹¹ Bonjour's promotion of a narrow view of mental content is important for the following reason. If non-experiential beliefs had their content in virtue of relations to external entities, then he would be unable to argue that empiricism can't justify principles of inference. For if beliefs do in fact have wide (relationally derived) content, then there would be no barrier (at least in principle) to the radical empiricist holding that general principles of inference can be derived from a series of experiential beliefs. For more on this, see Casullo, 2003, pp. 101-105.

⁹² Bonjour argues at length against representationalism and a Fodor-style LOT view of mental content in Chapter 6 of his *In Defense of Pure Reason*, and also in his (1991). However, I don't have the space here for a further look at his particular arguments, and will mainly set this issue aside. For a critique of these arguments, see Weller (1997).

⁹³ More specifically, the reasoning behind such a condition has to do with ruling out epistemic accidents: if a subject *S* supposedly has justification for believing a proposition *p*, yet there is no causal relation between the entity or region of reality in question and the belief *p*, *S*'s apparent justification for *p* would therein be undercut and defeated.⁹³

BonJour argues that such a causal condition is to be rejected in favor of an *instantiation* condition—that the universals don't have to be causally related to us/our mental events, or in other words that universals only have to be instantiated by us/our mental events. That is, it is a conceptual truth according to BonJour that when I think about an abstract concept like 'Greenness', the property *itself* is involved as an intrinsic and non-accidental constituent of the thought. The following lengthy passage makes this very clear:

If having a thought whose content is, for example, the claim that nothing can be red all over and green all over at the same time involves being in a mental state that instantiates the complex universal of which the universals redness and greenness are *literal constituents*, then at least much of the mystery surrounding my access to those universals and my ability to intuitively apprehend the relation of incompatibility between them is removed. In particular, there is no need to regard the apprehension of properties as a perceptual relation involving some mental analogue of vision...

[That is], the fact that I am in a state that instantiates a universal in which a particular property is a constituent ... *eliminates* the main metaphysical gap between mind and Platonistically construed universal.⁹⁴

The problem of epistemic access disappears, according to BonJour, when we understand the nature of thought: there simply is no gap between the acquainted-with⁹⁵ abstract property and mental instantiation of it (therein eliminating the spectre of epistemic accident). Accordingly, BonJour believes that the twin theses of Platonism and narrow mental content, i.e.:

(AR): Abstract entities exist, and they are instantiated by our concrete mental acts/events.

(NC): Thoughts (in BonJour's technical sense) involve or exemplify their content in a narrow, non-contingent manner. Thus, the content of a thought is both intrinsic and non-relational.

Entail a deflationary solution to the problem of epistemic access:

⁹⁴ BonJour, 1998, pg. 185. Italics added for emphasis.

⁹⁵ While he doesn't explicitly indicate so, BonJour's account of "direct" instantiation bares great affinity with Russell's theory of "direct acquaintance". For more, see Russell, (1912).

(EA): Because thoughts about abstract entities involve abstract entities as (intrinsic) constituents, there is no “access gap” between property and thinker.⁹⁶

3.3.2 *A Priori Justification*

With BonJour’s Platonism and its associated doctrines now briefly summarized, we can now proceed with an overview of his theory of a priori justification. Let us begin with BonJour’s preferred stock example of a belief justified a priori:

(RG): “Nothing can be red all over and green all over at the same time.”⁹⁷

And a (rather verbose) passage where he discusses this example:

First, I *understand* the proposition in question. This means that I comprehend or grasp the property indicated by the word ‘red’ and also that indicated by the word ‘green’, that I have adequate conceptions of redness and greenness ... Similarly, I understand the relation of incompatibility or exclusion that is conveyed by the rest of the words in the verbal formulation of the proposition, together with the way in which this relation is predicated of the two properties by the syntax of the sentence.

Second, given this understanding of the ingredients of the proposition, I am able to see or grasp or apprehend in a seemingly direct and unmediated way that the claim in question cannot fail to be true – that the natures of redness and greenness are such as to preclude their being jointly realized. It is this direct insight into the necessity of the claim that seems, at least *prima facie*, to justify my accepting it as true.⁹⁸

While his discussion is unclear, what BonJour seems to be suggesting, I think, is this.

My justification for believing the truth of (RG) is simply *a function of the very act of grasping or apprehending the abstract universals involved in the relevant thought* (plus the associated syntax). That is, he is suggesting that the relevant “phenomenological episode” (i.e., the “rational insight”) involves the direct instantiation of three universals: a relational universal (negation) and two colour universals (redness, greenness); where,

⁹⁶ BonJour, 1998, pp. 184-185.

⁹⁷ BonJour, 1998, pp. 100-101.

⁹⁸ BonJour, 1998, pg. 101.

assuming I understand (RG), I am therein disposed to “just see” that anything which instantiates one of the colour universals by necessity excludes the other.⁹⁹

In sum, BonJour is advocating a type of internalist foundationalism about a priori justification: in his view it is the *sui generis* act of rationally “seeing” or “grasping” necessity (via direct acquaintance with Platonic entities) which confers a proposition with *prima facie* justification. Furthermore, note that BonJour’s account is thus a variant of what I have designated ‘Extra-Mentalism’, at least insofar as the acts of rational insight, and the conceptual analyses we justify by appeal to such acts, are about an external (Platonic) reality.

3.4 Generality Again & the Methodological Upshot

With BonJour’s positive account of a priori justification (and its associated doctrines) summarized, we are now in a good position to both see how BonJour completes the Generality subargument and prevents its sceptical consequences from generalizing to a priori justification, evaluate the overall methodological upshot of his Deficiency argument against naturalism, and comment on the subsequent dialectical situation vis-à-vis my metaepistemological taxonomy.

3.4.1 Completing the Generality Subargument

⁹⁹ We can also see here, I think, how this account comports with BonJour’s dual definition of a priority on pg. 7 of his 1998: “There are two distinguishable aspects to ..., a priori justification, one negative, one positive: a proposition is justified a priori if it is justified [negative]: independently of any experience and [positive]” by appeal to reason or pure thought alone.”

In section 3.2.2, I indicated that without a plausible account of how non-experiential cognitive states), unlike experiential states, could have general content, Bonjour would be unable to prevent the key premise of the Generality subargument, viz.:

- (2). No cognitive state can directly justify a belief whose content goes beyond that of the state.

from generalizing to a priori justification (and henceforth rendering his position vulnerable to the same sceptical concerns motivating the argument in the first place).

We can now see, however, that Bonjour attempts to cash out the relevant distinction via the mechanics of his rationalism: unlike their experiential counterparts, non-experiential cognitive states have general content in virtue of directly (i.e. non-contingently) instantiating (certain sorts of) abstract entities. That is, on his account the tokening of the right sort of Platonic universal (e.g., necessity) in thought is simply a particular instancing of a general property, where the property itself is instantiated in virtue of being an intrinsic constituent of the relevant thought token (“cognitive state”).

Thus according to Bonjour, premise (2) doesn’t generalize for the following reason: whereas the content of experiential belief is limited to spatiotemporal particulars, non-experiential beliefs contain general abstracta as intrinsic content (including, importantly, logico-inferential principles).

3.4.2 The Purported Methodological Upshot

Given that Bonjour’s Platonism successfully provides the conceptual resources he needs to disarm a worry about the Generality subargument, it therein appears as if he can provide a cogent defence of (or more accurately, the key premise of) the Deficiency argument. Granting the soundness of this defense for the moment, what I would like to

do here is summarize and evaluate the purported methodological upshot of the Deficiency argument, and contextualize it vis-à-vis the broader contours of my discussion.

Let us first recall the elimination argument as I presented it above:

The Deficiency Argument:

- (1). Either (Radical) Naturalism is a minimally adequate meta-epistemological theory or Non-Naturalism is true.
- (2). If Naturalism is to be minimally adequate, then it mustn't imply scepticism
- (3). Naturalism implies scepticism.
- (4). Therefore Naturalism isn't minimally adequate.
- (5). Therefore Non-Naturalism is true.

BonJour sets the dialectic up as between a Quinean-style radical Naturalism (viz. "empiricism"), which holds that all justification ultimately derives from sensory experience, and his own Non-Naturalism, which admits of a robust faculty of Platonic rational insight. He then argues on sceptical grounds (via premise (3) and the supporting Generality subargument) that we must accept the latter:

[My central thesis is that] a rationalist [i.e. a Non-Naturalist] view of at least the moderate [defeasable] sort... is the only hope for a non-sceptical account of a priori justification and knowledge, and indeed for a non-sceptical account of knowledge generally (with the possible exception of those parts of empirical knowledge, if any, that can be fully justified by appeal to direct experience or observation alone).¹⁰⁰

I want to emphasize just how strong BonJour's conclusions are here: not only do sceptical considerations mean that empiricism is to be rejected; but rather such considerations imply that we *must* adopt a rationalist view of a priori justification (and as a consequence the Platonic Extra-Mentalism that accompanies it).

¹⁰⁰ BonJour, 1998, pg. 19.

Turning our attention now to my basic taxonomy outlined in the introduction, Bonjour has thus argued, by elimination, that the adoption of (ii) and (iii) is necessary to avoid scepticism, and furthermore – I emphasize – that adopting the doctrines in question *entails* a rejection of naturalism. Why? This consequence follows from Bonjour’s construction of the relevant metaepistemological dialectic; in his view, we *only* have two genuine choices¹⁰¹: (a) either we adopt a rationalist conception of a priority (where we grasp or apprehend a Platonic and extra-mental target domain of abstract entities), or (b) we adopt a Quinean empiricism, and therein capitulate to a global scepticism.

· 3.5 *Attacking Bonjour’s Non-Naturalism*

Now that I have provided an overview of both the negative and positive elements of Bonjour’s Non-Naturalism, I will now proceed to argue against it. More specifically, I will argue (through a series of steps) that premise (3) of the Deficiency argument is unsupported because question-begging: Bonjour has not successfully ruled out all available naturalistic alternatives with regards to a priori justification. Thus, we are not, as he maintains, forced by elimination to adopt his Platonic conception of rational insight in order to avoid scepticism, and as a result the argument fails.

3.5.1 *Setting Up*

¹⁰¹ To review, Bonjour thinks he can construe the metaepistemological situation in the manner he does because in his view, all “moderate” types of empiricism (i.e. those which countenance a linguistic or “analytic” conception of a priority) covertly collapse either into radical empiricism or rationalism. See Bonjour, chapter 2.

In order for my argument to proceed with clarity it will, I think, be necessary to take a more detailed look at how BonJour motivates the Deficiency argument. This will make it easier to see where exactly it goes astray, and provide a route for attack.

As I attempted to make clear above (3.4.2), BonJour's argument construes the relevant metaepistemological situation as one in which we are compelled to accept a rationalist account of a priori justification on grounds of the deficiency of naturalism. More specifically, his position is that naturalism either implicitly relies upon rational insight (in which case it corroborates his thesis), or it succumbs to a near total scepticism. In his own words:

...empiricist positions on a priori justification and knowledge ... are epistemological dead ends: the moderate empiricist attempt to reconcile a priori justification with empiricism by invoking the concept of analyticity does not succeed, indeed does not really get off the ground [i.e. it collapses into moderate rationalism]; and the radical empiricist attempt to dispense entirely with such justification ends in a nearly total scepticism. The indicated conclusion is that a viable non-sceptical epistemology rather than downgrading or rejecting a priori insight, must accept it more or less at face value as a genuine source of epistemic justification and knowledge. This is the main thesis of epistemological rationalism and also the central thesis of the present book.¹⁰²

Thus, according to BonJour, given the total inadequacy of empiricism (naturalism) to provide a viable and non-sceptical account of a priori justification, we arrive by elimination of the deficient positions to the conclusion that we *must* accept rationalism.

This negative argumentative strategy does, I note, have one clear advantage for BonJour: it allows him to dismiss the so-called "obscurity objection". This objection, as Paul Boghossian aptly puts it, claims that:

[T]he single most influential consideration against rational insight theories can be stated quite simply: no one has been able to explain— clearly enough—in what an act of rational insight could intelligibly consist. That is, no one has been able to say how some cognitive act, of a sort that we might plausibly enjoy, is able to yield immediate knowledge of the modal properties of properties. If the theory of rational insight is to serve as a genuine *explanation* for how we are able to have such a priori knowledge, rather than simply acting as a

¹⁰² BonJour, 1998, pg. 98.

placeholder for such an explanation, it must consist in more than a suggestive label; it must somehow lay bare, in appropriate detail, how some capacity that we have gets to work on the properties we are able to think about so as to disclose their natures.¹⁰³

In response, BonJour argues that any obscurity incurred by his rationalism is outweighed by the clear benefits of it being the only possible non-sceptical metaepistemology:

There is, to be sure, one reasonably clear sense in which many alleged rational or *a priori* insights are, if not necessarily *mysterious*, at least *inexplicable* in the sense of being apparently *irreducible*: they are apparently incapable of being reduced to or constituted out of some constellation of discursive steps or simpler cognitive elements of some other kind. But once it is realized that any such reduction [of a priori justification] would have to appeal to other apparent [rational] insights of a similar sort and thus ultimately, ... to irreducible ones ... it is hard to see why this admitted irreducibility should be thought to justify the charge of objectionable mysteriousness or lack of intelligibility.¹⁰⁴

Rational insight, BonJour admits, does countenance some irreducible and non-natural faculty by which we grasp abstract entities. However, he continues, once we realize that *any* account of a priori justification will have to make use of such a non-natural rational faculty in order to avoid scepticism the charge of obscurity no longer dialectically counts against rationalism.¹⁰⁵

In sum, then, BonJour's Deficiency argument stands or falls entirely on the soundness of its third premise, i.e.:

(3). Naturalism implies scepticism

for as we have seen, BonJour is making the case for his own metaepistemological position in an entirely negative manner. Hence, only if (3) and its supporting subargument successfully eliminates *all* non-sceptical naturalistic alternatives does this

¹⁰³ Paul Boghossian, "Inference and Insight", *Philosophy and Phenomenological Research*, 63 (3), 2001, pg. 636.

¹⁰⁴ BonJour, 1998, pg. 108.

¹⁰⁵ Unless, of course, we decide to take the sceptical route and simply deny both naturalism and rationalism. BonJour never directly addresses this position, however, and I shall follow suit (addressing scepticism would be another chapter altogether).

conclusion follow. In the next section, I will argue that Bonjour's doesn't, in fact, successfully rule out all the naturalistic alternatives.

3.5.2 Against the Deficiency Argument

With the preliminaries out of the way, I will now argue that the Deficiency argument fails. The reason for this failure, I will argue, is as follows: what I have identified as the key premise of the argument, i.e. (3), is question-begging; and as such the argument (at least as it stands) loses all justificatory support.

More specifically, because Bonjour attempts to justify the premise in question using the Generality subargument, my contention that (3) is question-begging will have to proceed through an engagement with the aforementioned subargument. As such, I will argue that even if we grant the Generality subargument, it is *too weak* to support the Bonjour's claim that naturalism implies scepticism; and as a result (3) is question-begging. And with (3) undermined, the Deficiency argument fails, because if naturalism doesn't imply scepticism then his conclusion by elimination (in favour of his Incompatibilist, Non-Naturalist version of rationalism) no longer follows

Why the Generality Subargument Doesn't Support Premise (3)

My first task is to demonstrate that the Generality subargument doesn't have the strong import (vis-à-vis his form of rationalism) that Bonjour assumes it does. To recall, the argument in question has the following form:

*The Generality Subargument*¹⁰⁶:

¹⁰⁶ I adopt this reconstruction from Casullo, 2003, pp. 103-104.

- (1). Experience is limited to *particular* objects.
- (2). No cognitive state can directly justify a belief whose content goes beyond that of the state.
- (3). Principles of inference are *general*.
- (4). Therefore, experience cannot directly justify principles of inference.

Assuming the argument is sound, (1)-(4) clearly establish a so-called narrow content¹⁰⁷ view of justificatory principles. For if (2) is correct and a cognitive state can't justify any belief whose content extends beyond the content of the particular state, then justification becomes a function of the belief's intrinsic (i.e. non-relational) properties.

The narrow view, however – and this I want to emphasize – is a doctrine which is *decidedly neutral* on the further issue of the ontology of concept possession. That is, it takes no determinate stance on the ontological status of the content of any particular mental state. For instance, while it is true that BonJour combines a narrow view of content with a Platonism about the ontology of concepts, this is, I wish to emphasize, *by no means necessary*. As such, the doctrine of narrow content endorsed above, i.e.:

(NC): Thoughts (in BonJour's technical sense) involve or exemplify their content in a narrow, non-contingent manner. Thus, the content of a thought is both intrinsic and non-relational.

is both *logically and conceptually independent* from the metaphysical doctrine of abstract realism also endorsed by BonJour, i.e.:

(AR): Abstract entities exist, and they are instantiated by concrete mental acts/events.

¹⁰⁷ The *narrow* content of a particular belief is a content of that belief that is completely determined by the subject's intrinsic properties (where an *intrinsic* property of an individual is a property that does not depend at all on the individual's environment. For example, having a certain shape is, arguably, an intrinsic property of a particular penny). For more, see Curtis Brown, "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/content-narrow/>.

Given this independence, then, there seems to be nothing *prima facie* stopping one from combining a narrow view of content like (NC) with a putatively naturalistic ontology of concept possession such as we see in, for example Block's 'Conceptual Role' semantics¹⁰⁸ or the Mentalist account endorsed by Goldman (assuming here that such a view is in fact viable, a point I shall be arguing for in Chapter 4).¹⁰⁹

If I am correct here, and (NC) is (at least *prima facie*) compatible with various ontological views of belief content, then the Generality argument is, we can contend, in and of itself *too weak* to establish premise (3) of the Deficiency argument. Why? Because if it is logically and conceptually possible to combine a narrow account of content such as (NC) with a Conceptual role or Mentalist ontology of concept possession, then Generality *per se.* doesn't imply that naturalism entails a radical form of scepticism. For this much stronger conclusion only follows if Bonjour is able to demonstrate *both of*:

(a) Narrow content is a necessary condition for the justification of principles of inference.

(b) No naturalistic ontology of concept possession is compatible with a narrow content view for such principles (because they must be constructed from experiential sources).

That is, premise (3) of the Deficiency argument only follows from (1)-(4) of the generality subargument if the following suppressed conditional premise is added to that subargument:

(5) If experience can't directly justify principles of inference, then nothing else available to the naturalist can (i.e. only rational insight can justify principles of inference).

¹⁰⁸ See for instance, Ned Block, "Advertisements for a Semantics for Psychology", *Midwest Studies in Philosophy* 10 (1):615-78. (1986).

¹⁰⁹ Very roughly, Mentalism proposes a view according to which a priori analyses and the application intuitions on which they are based, are about non-abstract, mental or psychological phenomena — concepts in a non-Fregean, non-Platonic sense. I don't actually endorse combining Mentalism with a narrow view of belief content; my point here (that they are compatible) being strictly argumentative. For more on Mentalism, see my Chapter 4.

For (1)-(4) plus (5) would imply the conclusion that naturalism entails scepticism; that is, if non-experiential narrow content is necessary for the justification of inference and the naturalist can rely upon or invoke *nothing else* other than experiential sources from which to justify such principles, then we get both (a) and (b). However – and this is the absolutely crucial point – premise (5) *begs the question*; for given the resources available to the naturalist with regards to combining a Mentalism (for example) about the a priori with a narrow view of mental content, Bonjour can't simply assume that experience is the only way the naturalist could justify principles of inference – for this would be to invoke as a premise the very point which is supposed to be at issue here.

Thus, I conclude that the Generality subargument is either too weak to justify premise (3), or it does so via an illicit and unsound suppressed premise. Either way, then, the crucial premise of the Deficiency argument lacks motivation.

Why Failure of (3) Undermines The Deficiency Argument

If my argument in the previous section is sound, and Bonjour's Generality subargument is question-begging, I am now in a position to step back and explain how the Deficiency argument fails.

At base, the reason the Deficiency argument fails has to do with the failure of (3) as its key premise. For if I am correct and (3) lacks justificatory support, then it is abundantly clear that Bonjour isn't warranted in jumping to his negative conclusion that we *must*, by elimination, accept his rationalist (and Incompatibilist) account of a priori justification. Why? Because I have shown that Bonjour hasn't made a sound case for the elimination of all naturalistic alternatives, and thus it is simply not true that we are

dialectically forced into adopting his rationalist account of the a priori (or an account of a priori analysis built on such an account) in order to avoid a radical scepticism. At most, Bonjour's argument, if sound, could establish the provisional claim that *if* a Quinean radical empiricism is the only viable naturalistic epistemology, then scepticism follows from naturalism (because such a radical empiricism fails to countenance the a priori at all). But this is, I must repeat, far short of the conclusion that naturalism *per se* implies a radical form of scepticism, which is *precisely what* he needs to establish in order to make a plausible case for the crucial premise (3) of his Deficiency argument.

What makes Bonjour's Deficiency argument look *prima facie* plausible, I think, is its rather stark presentation of the metaepistemological situation as involving a choice between a radical empiricism (which holds that all justified beliefs must ultimately rely upon experiential sources), and a non-naturalism involving the *sui generis* "grasping" of Platonic universals. However, as I have attempted to make clear, the situation is significantly more complicated; there are putatively naturalistic accounts of a priori justification, like Mentalism, which *don't* limit the justification of general principles of inference to experiential sources. Furthermore, overlooking or otherwise ignoring such naturalistic positions is, I think, a rather egregious error on Bonjour's part, one which ultimately undermines the soundness of his Deficiency argument.

3.6 Conclusion

With my objection to Bonjour's Deficiency argument complete, we can now assess the import of what has been discussed in this chapter with respect to both the

metaepistemological taxonomy introduced in Chapter 1, and my overall argumentative aims. There are, I think, two particular implications of importance for my purposes here.

First, I believe that we have good reason to reject Bonjour's case for the (*ICA*) principle. In my view, Bonjour has failed to provide a sound argument in defense of his primary thesis, i.e. that adopting (ii) and (iii) rationally requires a rejection of (i). That is, in his view, a Non-Naturalist, Incompatibilist form of rationalism is indispensable in order to build a viable account of a priori justification. However, if my objection to his Deficiency argument is sound, there isn't any clear barrier (at least in principle) to combining (i) with (ii)-(iii) in a way that Bonjour rejects. Thus, I argue, we have *no justified reason* to think that countenancing an account of a priori justification commits us to some type of Non-Naturalist and Incompatibilist form of rationalism; Bonjour's position appears to be motivated not by any metaepistemological realities pertaining to the a priori, but rather by his own previous commitments to such a form of rationalism. The bottom line, then, is that we are warranted in rejecting Bonjour's case for (*ICA*): we simply have no good reason to think that the adoption of (ii)-(iii) rationally requires acceptance of $\neg$ (i) because of the indispensability of rational insight.

Second, given that Bonjour explicitly structures the case for rationalism on indispensability (i.e. negatively) rather than fulfillment of any positive theoretical desiderata (for example, cf. the quote in 3.5.1), the failure of the Deficiency argument provides us with the foundation for a further argument against Bonjour's account. For if it is the case that: (a) Non-Naturalist rationalism *isn't* indispensable for a priori justification (which I have argued at present), and (b) there is in fact a viable naturalistic alternative, then I argue that, *ceteris paribus*, (c) we are warranted in adopting the latter

position. Why? Because the naturalistic alternative will, presumably, avoid the highly obscure metaphysics involved in the process of apprehending universals countenanced by Bonjour's rationalism. Now obviously, this argument will go through only if I am able to show that a viable naturalistic theory of a priori justification (and thus, potentially, of a priori analysis) is available. I will attempt to flesh out this argument in full in the following chapter.

Chapter 4: Goldman

4.0 Review of the Negative Project

Having objected to both forms of Incompatibilism (Chapters 2-3), I will now proceed to an exposition of my own Compatibilist view of a priori justification and conceptual analysis. Before I do, however, I would like to pause and recap how the aforementioned objections tie into the negative side of my metaepistemological project. This will provide a framework through which my positive aims will be situated and defended in this chapter.

First and foremost, I have tried to show that neither of Kornblith or BonJour's attempt to establish Incompatibilism has been successful. That is, I have argued that we have no good reason to accept the exclusionary principle (*ICA*) (i.e. that acceptance of (i) rationally requires rejection of (ii)-(iii) or vice versa)¹¹⁰ assumed by both the naturalist (Kornblith) and non-naturalist (BonJour) versions of Incompatibilism. In the case of the former, Kornblith hasn't plausibly shown that a priori analysis is unsuited to a naturalistically acceptable epistemology; and in the case of the latter, BonJour hasn't plausibly shown that a non-sceptical account of the a priori requires "rational insights" involving Platonic abstracta. Thus, I believe that we have no reason to accept either respective attempt to establish the (*ICN*) principle.

In sum: through the negative project I have attempted, via an attack upon Incompatibilism, to open up the logical and argumentative space required for me to proceed with my positive project; i.e. the articulation of a Compatibilist (i.e. Moderate Naturalist) account of a priori justification and analysis. To this task I now turn.

4.1 Introduction: The Positive Project

¹¹⁰ For more see section 1.3 of the first chapter.

In what follows, I will attempt to defend Moderate Naturalism through a Mentalist account of a priori justification inspired by the work of Alvin Goldman. More specifically, I will argue in favour of the following thesis: Moderate Naturalism has significant advantages over its immediate dialectical competitors. For clarificatory purposes, note that because my version of Moderate Naturalism adopts a Mentalist picture of conceptual analysis, talk here of ‘Mentalism’ will be interchangeable with ‘Moderate Naturalism’ as its metaepistemic analogue.

My general line of argumentation will proceed as follows. I will first begin (4.2) by providing a basic outline of Mentalism. I will then situate Mentalism vis-a-vis both Kornblith’s Strong Naturalism and Bonjour’s Non-Naturalism (4.3), arguing in a series of steps that it is the superior metaepistemological position.

4.1.1 Some Conceptual Presuppositions

In order to provide an informative sketch of both the basic contours of Mentalism and the argument I will present in its favour; it will first be necessary to make a few remarks about some of the philosophical presuppositions informing Goldman’s account. Note here, however, that this shall be nothing more than a bare-bones review and as such I leave the details to the relevant literature which I cite.

Getting down to it, there are two conceptual presuppositions in Goldman’s theory which, for clarificatory purposes, need to be mentioned. First:

(I): *Process Reliabilism* – Very roughly, Reliabilism is an externalist¹¹¹ account of epistemic justification which claims that the justificatory status of any token belief is a function of the truth-conductivity of the process through which it originated. That is:

(PR): *S's belief that p is epistemically justified if and only if S 's belief that p was produced by a reliable process.*

(where reliability, on Goldman's formulation, is a tendency to yield true beliefs, which will be an external feature of the agent S because it will supervene on relations between S 's psychological/neurological processes and the environment). The justified processes in question further break down into two categories: those which do not take in other belief states as inputs, and are thus foundational ("unconditionally justified"); and those which do take in other belief states as inputs, thus conferring justification on outputs only if the input belief-states are justified ("conditionally justified").¹¹²

(II): *Concept Possession as Classificational Competence*: Following Christopher Peacocke¹¹³, Goldman assumes that concept¹¹⁴ possession is a function of classificational competence. Ignoring various technicalities and complications, the account states that it is constitutive of determinately possessing a given concept F that a subject S will be disposed to affirm beliefs that accord with F 's semantic content, and deny beliefs that do not accord with F 's content. Determinate possession of an

¹¹¹ Where externalism, for present purposes, can be described as a view of epistemic justification which specifies that at least one of the conditions required for conferral of justification to any proposition is outside the mind of — i.e. external to the mental state of — the agent.

¹¹² For more detail on Process Reliabilism, see Goldman (1979), (1986), (1992).

¹¹³ Christopher Peacocke, *A Study of Concepts*, Cambridge: MIT Press, 1992. See especially Chapter 1 and his discussion of the "Principle of Dependence" pp. 3-7.

¹¹⁴ Be aware that 'Concept' can denote either the mental states of instantiating representational content, or the content of the relevant representation. Here I shall use it in the latter sense unless explicitly indicated otherwise.

individual concept, then, is on this account simply the sum of the dispositions which accord with *F*'s content and that any possessor *S* would, if prompted, affirm.

4.2. Mentalism (I): An Overview

The basic claim made by a Goldman-style Mentalism is that a priori analysis is fundamentally about *mental* (i.e. psychologically based) phenomena. More specifically, when we assess application intuitions (i.e. non-inferential judgments about the applicability of a predicate or conceptual term to particular situations) in the exercise of conceptual analysis, we are activating our dispositions to apply a *personal, psychological concept* to the particular scenario. In this section, I will attempt to flesh out this approach in some detail.

First, however, I would like, for contextual purposes, to frame the discussion with a review of the basic account of conceptual analysis I outlined in the introductory chapter.

4.2.1 Conceptual Analysis: a Redux

In its most general guise, my basic account of conceptual analysis takes the consistency (or inconsistency) of application intuitions with a proposed conceptual analysis of a relevant concept or term as *prima facie* evidence for the truth (or falsity) of the analysis.

More specifically, I have maintained that conceptual analysis proceeds as a two-stage a priori exercise; which can be summarized as follows. First, for any general concept *F* and the set of modal (i.e. hypothetical as well as actual) scenarios $e_1...e_n$, subject *S* makes an initial assessment of her intuitions regarding the application (or non-

application) of F to $e_1...e_n$, where these intuitions are taken as a priori (but fallible and defeasible) expressions of the content of S 's concept F . Secondly, S then uses these application intuitions as her basis for arriving (inferentially) at an a priori justified belief in the truth (or falsity) of a proposed analysis of F .

4.2.2 Goldman's Basic Account

With the background in place, we can now turn to an examination of how a Mentalist account of the two-stage process outlined here would run.

As articulated above, the fundamental starting point for Goldman's position is this: concepts, construed as the targets of philosophical analyses, *are literally in the mind* (as mental representations or semantic categories which supervene upon psychological states). That is, when S assesses the content of her concept F , she is assessing application intuitions regarding the content of her personal, psychological concept of F . In other words, Mentalism is an explicitly *psychologistic and naturalized* account of conceptual analysis. As Goldman and Pust put it:

Mentalism interprets philosophical analysis as trying to shed light on the *concepts* behind philosophically interesting predicates, where the term 'concept' refers to a psychological structure or state that underpins a cognizer's deployment of a natural language predicate. Thus, Jones's concept of 'apple' is the psychological structure that underlies her deployment of the predicate 'apple', and Jones's concept of knowledge is the psychological structure that underlies her deployment of the predicate 'knows'... the structures in question presumably include some sorts of mental representations as well as some kinds of processing routines.¹¹⁵

More schematically, we can put it this way: according to Mentalism, determinate concept possession involves the dispositional ability to competently deploy or apply the relevant concept to hypothetical scenarios or situations; where evidence for the

¹¹⁵ Alvin Goldman and Joel Pust, 'Philosophical Theory and Intuitional Evidence'. In M. DePaul & W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield, pp. 187-188.

application (or non-application) of a concept ' F ' to a scenario e will take the form of an occurrent or spontaneous application intuition of the form "scenario e does (does not) satisfy the concept I express through the predicate ' F '".¹¹⁶ These application intuitions are then taken as basic evidence for an analysis of F .

What is most important to note about this account of conceptual analysis for my purposes is the following: according to Mentalism, both the relevant application intuitions, and the general concept they serve as evidence for are construed as being about mentally-based (i.e. psychological/neurological) phenomena. So Mentalism *doesn't* rely upon or commit to a view of analysis such that it involves the accessing of abstract or otherwise extramental entities of any kind.

Is a conceptual analysis really an a priori endeavour on this account? One might worry that psychologism renders conceptual analysis an empirical affair, and thus inconsistent with our Standard Philosophical Methodology (as I have construed it) because the entities it is about (concepts in the personal, psychological sense) are contingent, empirical phenomena. But, to respond to this worry, recall my opening characterization of a priori justification in Chapter 1. Given this characterization, application intuitions, and conceptual analyses, can be justified a priori, despite the fact that the entities they are about are empirical, "in the mind" phenomena.

Note, however, that this explanation is conditional upon the assumption that application intuitions (and the beliefs in conceptual analyses we inferentially base upon

¹¹⁶ Note that application intuitions are occurrent mental events — judgments about whether the relevant term/predicate applies. They are thus not the dispositional state that is supposed to be identified with possession of the concept that the term/predicate expresses, nor with the underlying psychological structure that is the concept. (To judge, occurrently, that e_1 is F , that e_2 is F , that e_3 is F , that e_4 is not F , amounts to having an application intuition; however none of the intuitions is identical to the personal concept expressed by ' F ', nor to the dispositional state that amounts to possessing this concept.) Thanks to David Matheson for clarifying this point.

them) are, on the Mentalist account, produced by reliable processes. I will argue in the next section in favour of the reliability of the relevant processes; but for my present expository purposes the conditional explanation is sufficient.

4.3 Mentalism (II): Arguing in its Favour

With a basic sketch of Mentalism on the table, I will now proceed to argue in favour of the two theses I outlined above. That is, I will argue that Mentalism (and Moderate Naturalism as its metaepistemological analogue) is superior to its primary metaepistemological rivals.

In order to set up my argument, however, I will first need to introduce a basic schema of evaluation, centered upon a set of plausible epistemic desiderata.

4.3.1 Evidentially Basic States

As I mentioned in Chapter 1, the application intuitions serving as data for conceptual analysis are taken by Standard Philosophical Methodology (hereafter SPM) to be evidentially basic states which provide justificatory warrant for accepting or rejecting analyses of concepts. But what are evidentially basic states, anyway? Here is how Goldman and Pust define them (and I shall herein adopt their construal):

Many classes of mental states comprise basic evidential sources, including visual seemings, auditory seemings, memory seemings, and perhaps introspective seemings. In each case, there is a class *M* of mental states such that being in one of these mental states is *prima facie* evidence for the truth of its content, at least when the (token) state occurs in favorable *M*-circumstances.¹¹⁷

¹¹⁷ Special thanks to David Matheson for his help here. Also, compare Alvin Goldman and Joel Pust, 1998, pg. 179-180.

That is, an evidentially basic state is a state such that being in it provides *prima facie*¹¹⁸ evidence for the truth of its content. So, for instance, given that visual seemings are evidentially basic, my seeming to see that a yellow pencil is on the desk comprises *prima facie* evidence for the truth of the proposition that there is a yellow pencil on the desk.

The following is, I think, a very plausible condition on any state's being evidentially basic in this sense:

(*RI*): Mental states of type *M* are evidentially basic only if *M* states are reliable indicators of the truth of their contents (or the truth of closely related contents), at least when the *M* states occur in *M* favorable circumstances.¹¹⁹

Where being a reliable indicator in turn requires two things, as Goldman and Pust note:

(a). *Presence of a Causal Medium*: there must be a distinctive causal relationship or route between the family of state of affairs which make *M* states true or false and the *M* states themselves. For instance, in the case of visual seemings, work in cognitive science has given us a very good explanation of the causal route that obtains between light reflecting off an external state of affairs, the perceivers' retinas, and information travelling (through the visual cortex) to the brain, producing a visual seeming of the relevant state of affairs.¹²⁰

(b). *Counterfactual Dependence*: there must be a counterfactual dependence of the relevant *M* states upon what makes them true. The notion of counterfactual dependence can be cashed out as follows: for any particular mental state *m* and fact *f* that makes its content true, we can say that *m* counterfactually depends on *f* just in case had *f* not obtained, *m* would not have obtained. So, for example, returning to evidentially basic

¹¹⁸ *Prima Facie* as in the absence of defeaters.

¹¹⁹ Goldman and Pust, 1998, pg. 180.

¹²⁰ Ibid, pg. 180.

sources of belief, we can say that my visual experience as of a yellow object in front of me counterfactually depends on the fact that there is a yellow object in my visual field, in the sense that, under favorable environmental conditions, had there been no yellow object in my visual field, I would not have had a visual experience as of there being a yellow object before me, etc.¹²¹

4.3.2 The Argument from Explanatory Comprehensiveness

Now the fundamental question that arises here, I think, is this: are application intuitions evidentially basic in the foregoing sense? Do they satisfy the Presence of a Causal Medium and Counterfactual Dependence conditions? Different interpretations of what these intuitions are about (i.e. different interpretations of the *target domain* of the intuitions) make the chances of satisfying these conditions better or worse; and in what follows, I will argue that only Mentalism – as opposed to Non-Naturalism or Strong Naturalism – provides an adequate understanding of how application intuitions can successfully satisfy these conditions, and thus provide an cogent account of what we are doing in using application intuitions and conceptual analysis as part of our SPM.¹²²

Mentalism vs. Extra-Mentalism

The first thing to note is that Mentalism, which construes the target domain of application intuitions (as well as the target domain of the conceptual analyses they are used to

¹²¹Ibid, pg. 180. For more, see David Lewis (1986).

¹²² Note that this discussion draws heavily on the discussion of similar issues in Goldman and Pust, 1998, pp. 178-188. My contribution is to transpose the distinctions they make into my discussion of the relevant metaepistemological positions at stake. And while I don't have the room to argue for and justify the two conditions operative in (RI), note that they are not particularly controversial, and as such I don't think that it would be hard to provide the relevant justification. Goldman (1999) and (1986) discusses the reliability of evidentially basic M states at length, and I direct the interested reader to these for more.

justify) as consisting of “in the mind” phenomena, can be contrasted with both Bonjour’s Non-Naturalism and Kornblith’s Strong Naturalism in the following manner: unlike the former, both of the latter positions construe the target domain of application intuitions as consisting of *external, extra mental phenomena*. In other words, both of the latter agree inasmuch as they both take the targets of epistemological inquiry to be “outside of the mind” – i.e. for Kornblith, the relevant targets are natural kinds; for Bonjour, they are abstract, Platonic entities. Thus, on this view, our application intuitions provide an evidential route to (or elucidate the objective content of) the properties of some natural kind or abstract entity (which our concept *F* represents or instantiates).

So, in this section, and the remainder, of the chapter, when I refer to ‘Extra-Mentalism’, I shall be referring to a general term which denotes any position (including those of Bonjour and Kornblith) which construes the target domain of philosophical application intuitions as consisting of external, “free-floating” entities.

(A) Extra-Mentalism and (RI)

With the Mentalism/Extra-Mentalism dichotomy outlined, I can now proceed to the first point I wish to make in this section, which is this: if the target domain is construed in an Extra-Mentalist manner, application intuitions cannot satisfy the two conditions for being reliable indicators – the Presence of a Causal Medium condition, and the Counterfactual Dependence condition. In other words, both types of Extra-Mentalism (i.e. Non-Naturalism and Strong Naturalism) present a picture of application intuitions such that these intuitions would fail the conditions in question. Let us consider each type of Extra-Mentalism in turn.

Non-Naturalist Extramentalism (BonJour): Rationalism, of the kind advocated by BonJour, assumes that the evidential route from application intuition to abstract entity obtains in some type of spontaneous episode of “rational insight” or “apprehension”. That is, it invokes a quasi-perceptual model of the medium through which abstracta is grasped.¹²³

Now the problem with this view should be obvious: *how could* an application intuition, as a dated, *mental* episode, be a reliable indicator of the properties of abstract and timeless universals? That is, how could our intuitions provide evidence for the truth of some non-spatiotemporal state of affairs? It seems unlikely that they could. Furthermore, the talk of “apprehension” is left entirely metaphorical, and is thus unhelpful; for unlike genuine perceptual processes like vision, where we have a good grasp of the basic causal mechanisms, the supposed ability to apprehend entities outside of space and time is totally mysterious. As a result, the truth-conditions for such claims are totally opaque and defeasable. As Goldman and Pust put it:

Is there any reason to believe that intuitions could ever be reliable indicators of a universal’s positive and negative instances (even under favourable circumstances)? The problem is the apparent “distance” or “remoteness” between intuitions, which are dated mental states, and a nonphysical, extra-mental, extra-temporal entity. How could the former be reliable indicators of the latter?¹²⁴

I contend that there is no way for application intuitions to fulfill the Presence of a Causal Medium condition or the Counterfactual Dependence condition on the Non-Naturalist version of Extra-Mentalism. That is, I can see neither (a) any cogent causal story about how we could reliably grasp abstract entities, nor can I see how, (b) even if

¹²³ For instance, see BonJour, *In Defense of Pure Reason*, pp. 180-186 and Plantinga (1993).

¹²⁴ Goldman and Pust, 1998, 185.

we did happen to accurately grasp them, how this grasp could be counterfactually dependent on the relevant abstracta, in the absence of a cogent causal relationship.¹²⁵

Strong Naturalist Extra-Mentalism (Kornblith): What about Kornblith's "Natural Kinds" approach to epistemological method; is it a more successful form of Extra-Mentalism, with respect to (a) and (b)? I argue that it is not.

The main problem with this approach, in my view, is very similar to the problem encountered by Bonjour's approach. That is, it seems doubtful if application intuitions could provide a reliable evidential route to external natural kinds; i.e.:

The problem is: how can intuition get any reliable purchase on the constitutions of a natural kind and an example so as to decide whether the latter matches the former? ... If the judgements are about natural kinds and their putative instances, a cognizer must first discern the underlying nature of the natural kind, then discern the underlying nature of the example, and then decide whether these match. It is not clear how intuition could contribute to the reliable execution of these tasks.¹²⁶

The important point here is this: our application intuitions can only get purchase on the a priori content of conceptual terms. However, if the relevant term has a natural kind as its content, as is the case with terms like 'Gold' or 'Water', then how are our putatively a priori intuitions about the application of the term going to reliably elucidate the relevant kind? I don't see how they could. Furthermore, although it is remotely possible that we could somehow latch on to the right properties of the kind a priori, this would be a case of epistemic luck, and hardly such as to make our application intuitions capable of satisfying the Presence of a Causal Medium and Counterfactual Dependence conditions. So, the bottom line here is that I see neither how (a) a genuine causal route

¹²⁵ More specifically, my view is that the lack of a causal medium suggests that if we did happen to grasp the properties of some universal without a cogent causal story, the grasp would be unreliable; thus falling the counterfactual condition. This line of argument is similar to that of Benacerraf (1973). See Paul Benacerraf, "Mathematical Truth", *Journal of Philosophy* 70 (19), 1973, pp. 661-679.

¹²⁶ Goldman and Pust, 1998, 187.

could obtain between the intuitions and the content (understood as a natural kind), nor (b), how a reliable counterfactual relationship could obtain between the two (given the lack of a plausible causal story).

As a result, I believe that Kornblith's account has no way of successfully accounting for how application intuitions, as appealed to in SPM, could have *any evidential weight* vis-a-vis the specified target domain.¹²⁷ He might object, in the end, that his preferred method is one that explicitly *rejects* appeal to application intuition, so my argument here simply misses the target. However if my Starting Points objection is correct, he is committed to relying upon application intuition nonetheless. As such, then, I don't think that Kornblith's metaepistemology can avoid facing criticisms of the sort articulated at present.

To sum up the above: If I am correct, then neither Bonjour nor Kornblith's metaepistemology meets the two conditions required for application intuitions to count as reliable indicators. Thus, on neither account does it seem plausible that application intuitions could serve as evidentially basic states.

So, neither of the relevant authors can provide a cogent description of what is going on in our SPM; for they would each be committed to holding that the general method employed by philosophers in the attempt to elucidate the structure of philosophical concepts is unreliable and therefore completely unjustified vis-a-vis the

¹²⁷ Kornblith could here protest that I am being unfair, inasmuch as his position deliberately discards much of Standard Philosophical methodology as irrevocably flawed. I don't deny that he can make this move, but note that the issue I am concerned with at present regards the justification of a priori analysis *as it is standardly construed*. Thus, even if he wants to discard standard methodology, this doesn't in any way preclude my point that his method has no way of accounting for it. Furthermore, as I argued in Chapter 2, we have good reason, independent of any consideration of SPM or (RI), to reject his approach.

relevant target domain. And this, I think, is a strong *prima facie* reason to doubt both Non-Naturalism and Strong Naturalism; at least in the absence of a very convincing argument to the effect that the majority philosophers have been up to this point relying upon fundamentally flawed methodology.

(B) Mentalism and (RI)

I want to now turn to the second point I want to make in this section, which is that Mentalism – in contrast to Extra-Mentalism, *can*, in fact, allow us to see application intuitions as satisfying the two conditions associated with (RI). This will set up my defense of the thesis I outlined at the beginning of this chapter.

How does a Goldman-style Mentalism satisfy the two conditions in question? To demonstrate this clearly, I will consider Mentalism vis-a-vis each condition in turn.

(a) Mentalism as satisfying the Presence of a Causal Medium Condition: The causal condition, to review, states that there should be (evidence for) the existence of a distinctive causal relationship or route between the family of state of affairs which make the relevant reliable indicator states true or false and being in those states themselves. So, the question here is this: is there evidence for a distinct causal route between concepts, construed as psychological entities, and application intuitions about the terms/predicates expressing those concepts?

I think that there is. The reason is rooted in the fact that unlike the Extra-Mental views surveyed above, Mentalism is not burdened by attempting to show how application intuitions, as dated mental events, could somehow be evidence for the presence or absence of properties of an abstract entity or natural kind. For the semantic

content(s) underpinning concept deployment, on the Mentalist account, are embedded in, and supervene upon, *psychological states/events* (which in turn are embedded in neurological processes). Thus, it is comparatively easy to see how a causal route could be borne out between concepts structures and their conscious manifestations in application intuitions.¹²⁸

Though we don't know the exact causal structures leading from concepts (and their possession) to their corresponding application intuitions, I admit, there is much evidence for the existence of such a causal route provided by cognitive science¹²⁹; and even at this early stage in the development of such theories, it is far less obscure than anything offered by the Extra-Mentalist. Furthermore, in anticipation of potential objections from those worried about the lack of a precise understanding of the causal route vis-a-vis reliability, note that we haven't worked out the precise structures responsible for the production of visual, memorial or auditory experiences or seemings as basic evidential states; yet no one (on pain of a general scepticism)¹³⁰ seriously doubts that such processes are sufficiently reliable. So, I don't think that such an objection would hold any traction.

(b). *Mentalism as satisfying the Counterfactual Condition*: I also think that Mentalism can allow us to see how application intuitions can satisfy the counterfactual condition on reliable indication as specified by (RI). That is, there is reason to think that the Mentalist account of the targets of application intuitions can support the existence of a

¹²⁸ See Goldman and Pust, 1998, pp. 180-183.

¹²⁹ For one example, see David Marr's (1982) pioneering work on the computational theory of vision.

¹³⁰ George Bealer argued for this in his '*The Incoherence of Empiricism*'. *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol.66, 1987, pp. 99-143.

robust counterfactual relationship between these targets and their corresponding application intuitions.

The reason I think that Mentalism can, in fact, support the relevant counterfactual relationship has to do with the account of concept possession as rooted in classificational competence. As Goldman puts it:

... it is easy to see that intuitions ... can be reliable indicators of the truth of their associated [contents]. That is because there might well be a counterfactual dependence between x 's satisfaction or non-satisfaction of the concept expressed through F and what the intuition "says" about x ... a concept tends to be manifested by intuitions that reflect or express its content.¹³¹

In other words, the satisfaction of a concept F by a particular e in a given scenario is a function of subject S 's classificational competence in the deployment of F ; given that it is constituent of possessing a psychologically embedded concept that S will be disposed to affirm positive instances of F and deny negative instances, we can say that (assuming normal semantic and linguistic competency) the following counterfactual holds: had F not been satisfied by e , then S *would not* have affirmed this.¹³²

Is this account of classificational competence sufficiently robust, counterfactually speaking? I believe that it is. Of course, the disposition to correctly classify particulars falling under concepts isn't infallible, and there are circumstances which are likely to produce inaccurate classifications.¹³³ However, there is good reason

¹³¹ Goldman and Pust, 1998, pg. 189.

¹³² As Goldman puts it on pg 15 of his 2007: "It's part of the nature of concepts (in the personal psychological sense) that possessing a concept tends to give rise to beliefs and intuitions that accord with the contents of the concept. If the content of someone's concept F implies that F does (does not) apply to example x , then that person is disposed to intuit that F applies (doesn't apply) to x when the issue is raised in his mind." See '*Philosophical intuitions: their target, their source, and their epistemic status*'. Grazer Philosophische Studien, 2007. 74, pp. 1-26.

¹³³ More specifically, there are situations which are likely to distort classificational dispositions. The most important are situations of "theory contamination"; viz.. situations in which application intuitions are influenced by an explicit theory of the relevant concept, rather than the content of the concept itself. This explains why some philosophers place so much evidential weight upon "folk" intuitions – they are less

to think that, analogous to other basic evidential states, under appropriate circumstances the use of application intuitions by competent language users is sufficiently reliable to satisfy (RI).

To sum up: given what I have said above, I believe that Mentalism can satisfy both the Presence of a Causal Medium and the Counterfactual Dependence conditions associated with (RI), and thus can adequately explain how application intuitions, in line with our SPM, constitutes basic evidential sources for theories or analyses of concepts of philosophical interest. Furthermore, because we can explain (at least in outline) both the causal route and counterfactual dependence relation underpinning the deployment of a personal psychological concept in the evaluation of a hypothetical scenario, it is easy to see how they could be *prima facie* warranted.

Thus, it is clear that of the accounts currently under consideration, Mentalism is the more explanatorily acceptable vis-à-vis our Standard Philosophical Methodology. Why? It is the *only* account of the three which accommodates the very plausible assumption made by SPM that application intuitions constitute basic evidence for the truth of philosophical theories.

4.3.3 *The Argument from Parsimony*

My second argument for Moderate Naturalism pertains to considerations of parsimony (theoretical simplicity); and is thus independent of considerations regarding what I have called our SPM. That is, I argue that Moderate Naturalism (Mentalism) is more

likely to be contaminated by an explicit theory of the target phenomena. For more, see Goldman and Pust, 1998, pp. 188-189.

parsimonious (metaphysically speaking) that its Extra-Mentalist rivals, and this gives us reason to prefer the former to the latter.

As a point of departure, let's examine the question of how the various metaepistemologies in question construe the *target domain* of application intuitions. To recall, both BonJour and Kornblith (in different ways) construe of the target domain as consisting of entities "outside of the subject's mind"; whereas a Goldman-style Mentalism construes relevant target domain as consisting of "inside the mind" entities. In other words:

<p style="text-align: center;"><i>Extra-Mentalism</i></p> <p><i>Strong Naturalism</i> (Kornblith): The target domain of application intuitions consists of extramental natural kinds.</p> <p><i>Non-Naturalism</i> (BonJour): The target domain of application intuitions consists of extramental abstract entities.</p> <p style="text-align: center;"><i>Mentalism</i></p> <p><i>Moderate Naturalism</i> (Goldman): The target domain of application intuitions consists of psychologically-based concepts.</p>

As we have seen, understanding how the metaepistemological positions in question differ on the relevant target domain is important with regards to understanding how each fits (or does not fit) with our SPM.

That being the case, the question of the target domain of application intuition also, I think, has significant bearing on the *metaphysical* status and commitments of each position; commitments which, I think, contribute to the overall plausibility of the relevant position. More specifically, if we believe that parsimony is (when paired with

plausibility)¹³⁴ a desirable constraint upon theory construction, we can evaluate each position vis-a-vis the metaphysical parsimony of its particular construal of the target domain. I do believe that it is a desirable constraint or desiderata; and thus I argue that because Mentalism construes the target domain of application intuitions (and of the conceptual analyses they support) in a more parsimonious manner than its Extra-Mentalist rivals, it is *ceterus paribus* more plausible.

Now, on to the argument itself. First, note that I will cash out 'parsimony' in an epistemic sense; i.e. in terms of what is most rational or reasonable to believe:

(MP): Other things being equal, if T_1 is more metaphysically parsimonious than T_2 then it is rational to prefer T_1 to T_2 .¹³⁵

Where 'metaphysical parsimony', in particular, amounts to roughly the following. A theory, T , is metaphysically committed to F s if and only if T entails that F s exist. If two theories, T_1 and T_2 , have the same metaphysical commitments except that T_2 is metaphysically committed to F s and T_1 is not, then T_1 is more parsimonious than T_2 .¹³⁶

Why do I say that Mentalism is more parsimonious than its competitors? The easiest way to see this is to contrast Mentalism to each form of Extra-Mentalism considered here:

(A) *Mentalism vs. Non-Naturalism*: Here it is easy to see that the latter is more parsimonious than the former. The reason is as follows: Mentalism construes the target domain of application intuition as consisting of personal, psychological concepts,

¹³⁴ That is, we have to balance parsimony with plausibility: an ontology with nothing in it would obviously be more parsimonious than an ontology containing various entities, but it would obviously be highly implausible. When I talk about parsimony, then, I am assuming that considerations of parsimony are being applied to theories of (very roughly) comparable plausibility.

¹³⁵ Alan Baker, "Simplicity", The Stanford Encyclopedia of Philosophy (Summer 2011 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2011/entries/simplicity/>>.

¹³⁶ Ibid.

whereas Non-Naturalism construes it as consisting of Platonic universals. Thus, the latter is existentially committed to a type of entity that the former is not; and as a result we can say that Mentalism is more parsimonious than Non-Naturalism.

(B) *Mentalism vs. Strong Naturalism*: We also have reason to think that a Goldman-style Mentalism is more parsimonious than Kornblith's Strong Naturalism. Again, the reason has to do with construal of the target domain: Strong Naturalism, eschewing the role of the *a priori*, reconstructs epistemology as the empirical investigation of natural kinds. As a result, Strong Naturalism, in order to get its methodological program off the ground – is existentially committed to a class of *philosophically investigable* natural kinds; i.e. the kinds which play the role of truth-maker in an empirically-based epistemology – which Moderate Naturalism is not.¹³⁷ The point here is this: like Non-Naturalism, Strong Naturalism is thus committed to countenancing an extra kind of entity; and as a result, the latter appears to be *ceteris paribus* the more parsimonious, at least vis-a-vis (*MP*).

As such, it appears that Mentalism, insofar as it locates the relevant target phenomena as in the mind, is more metaphysically parsimonious than either Non-Naturalism or Strong Naturalism. And this, I think, gives us another *prima facie* reason to prefer Mentalism.

To summarize the overall upshot of section 4.3, I think that the argumentative prospects for a Goldman-style Mentalism regarding *a priori* analysis are very good, at least with

¹³⁷ Note however, that I am not saying that Moderate Naturalism isn't thus committed to the existence of a wide variety of natural kinds (it is). I am only saying that Strong Naturalism is existentially committed to an *extra* class of natural kinds – that is, *natural epistemological kinds* – which the former is not.

regards to the present taxonomy. Thus, my conclusion is that Moderate Naturalism is superior to both Bonjour's Non-Naturalism, and Kornblith's Strong Naturalism.

4.4 Objections and Replies

Given the controversial nature of the present discussion, I am aware that my account will be subject to numerous objections. Thus, in the final section of this chapter I would like to briefly single out what I think are the two most pressing objections to Mentalism, and attempt to sketch out the beginnings of a reply to them.

4.4.1 Objection # 1: "Isn't Mentalism too Subjective?"

A very reasonable worry about a Goldman-style mentalism about a priori analysis is that it is somehow too subjective, i.e. that it is relativistic or overly concerned about personal concepts; thus rendering the method trivial or at least of lesser importance than it is taken to be by our SPM. To put the objection differently, one may have thought that the analyses supported by our application intuitions are ultimately about objective matters of fact, not about our own psychological structures. Doesn't the Mentalist approach, in construing the target domain of application intuitions and the analyses they support analysis as consisting of personal, psychological structures, somehow make this impossible?

I think the answer is no. To see why, let me first clarify the basic objection as a means of singling out exactly where it goes wrong. I take it that the following is a plausible reconstruction:

The Subjectivity Objection:

1. If Mentalism is true, then a priori analyses are only ever about subjective entities (viz. about concepts as personal, psychological structures).
2. If a priori analyses are only ever about subjective entities, then a priori analyses could only ever provide insight into the particular mental lives of those who accept the analyses.
3. But a priori analyses can provide insight into things other than the particular mental lives of those who accept the analyses.
4. Therefore, Mentalism is not true.

My response is to deny Premise 2. That is, I think it is false to think that Mentalism (as a subjective account of a priori analysis) can only provide insight into the mental lives of those who accept the particular analyses. Why do I deny it?

My grounds are as follows: provided the way in which the subjective entities (concepts as personal, psychological structures) get into the minds of those who accept the analyses is sufficiently similar to the way in which analogous subjective entities get into the minds of others (i.e. provided there are shared means of concept acquisition; for example by social inculcation or by innate biological mechanisms), it is very likely that the subjective entities in the minds of others are going to be sufficiently similar to the subjective entities in the minds of those who accept the analyses to ensure that those analyses can provide insight into the mental lives of others *and not just into the particular mental lives of those who accept the theories.*¹³⁸ As Goldman puts it in a recent paper:

What, then, is the central type of hypothesis for which an intuition is evidence? The prototypical form of such a hypothesis will be “Case C meets the satisfaction conditions for F”, where F is the communal meaning of the queried predicate. This form of hypothesis is also the content of the subject’s intuition; it is what the person intuits.¹³⁹

In other words, the personal, psychological concepts deployed by a linguistically competent subject have socially or communally determined deployment or satisfaction

¹³⁸ Thanks to David Matheson for some helpful suggestions regarding this point.

¹³⁹ Alvin Goldman, ‘*Philosophical Naturalism and Intuitional Methodology*’. Proceedings and Addresses of the American Philosophical Association, 2010, pg. 16.

conditions (i.e. truth-conditions), implying that a priori analysis will provide insight into *shared* content.

So I don't think it at all follows from the fact that Mentalism construes concepts as personal and psychological that a priori analysis of concepts is somehow going to be wildly subjective or of only idiosyncratic interest – if it is the case that concepts have communal acquisition conditions, then the a priori analysis of concepts will be sufficiently objective on the Mentalist construal. And as a result, I believe that the Subjectivity Objection fails.¹⁴⁰

4.4.2 Objection #2: The Reliability of Intuitions

The second objection pertains to concerns about the putative reliability of our application intuitions. Many, especially those committed to the negative experimental philosophy program¹⁴¹, have expressed doubts about whether application intuitions constitute a basic and reliable source of evidence for the truth of analyses of the concept(s) in question. I will here briefly look at two more specific worries about reliability, and provide responses.

(1) Responding to the Source Opacity Objection

¹⁴⁰ Here is an excellent quote along the same lines courtesy of Paul Grice, from his *Studies In The Ways Of Words*, Cambridge, MA: Harvard University Press, 1989, pg. 175: “You may regard me, when engaged in a piece of conceptual analysis, as primarily concerned to provide a conceptual analysis of *my own* use of a given expression... to reach a satisfying conceptual analysis of one's own use of an expression is often extremely difficult, and you must expect most of my discussion about the conceptual analysis to relate to this difficulty. But if I think that I have reached a satisfactory conceptual analysis of my own use, I do not go on to conduct a poll to see if this analysis fits other people's use of the expression. For one thing, I assume (justifiably, I think) that it does in general fit other people's use, for the expression for which (as a philosopher) I am normally concerned are pretty commonly used ones; and if a particular expression E was given by some of the people with whom I talk in my daily life a substantially different meaning from the one in which I gave it, then I should almost certainly have discovered this; one does discover people's linguistic idiosyncrasies.”

¹⁴¹ Cf. Weinberg, Nichols and Stich (2001), Weinberg (2008).

As I alluded to above, some may be worried that the causal source of application intuition may be too opaque to confer reliability upon its outputs. That is, it may be objected that our understanding of the neuropsychological base of concept deployment is still quite obscure, and that this obscurity severely precludes its putative reliability.

To this I respond as follows. While it is true that we have yet to work out the details of the neuropsychology of concept deployment, the situation here is no different than for other purportedly basic sources of evidence such as memory and perception; insofar as we have yet to work out their precise neuropsychological causal pathways. However – this being the key point – on pain of general scepticism, we *don't* seriously doubt the existence and reliability of the relevant causal pathways. Thus, I argue, without an argument demonstrating a clear disanalogy between perception and memory as reliable sources of evidence and concept deployment (in terms of reliability), we have no reason to think that outputs of the latter are *prima facie* unreliable.

(2) Responding to the Relativism Objection

According to Weinberg, Nichols and Stich¹⁴², we have reason to think that application intuitions vary across cultural bounds. And they argue that this is a serious problem, at least insofar as SPM assumes that certain norms and concepts are invariant in content, i.e. that something along the lines of premise (3) of the Subjectivity Objection (given above) is true.

I have two brief responses to this objection:

¹⁴² Jonathan M. Weinberg, Shaun M. Nichols, and Stephen Stich. 'Normativity and Epistemic Intuitions'. *Philosophical Topics*, 29 (1-2), 2001, pp. 429-460.

(a). As Jennifer Nagel¹⁴³ has argued, we have reason to doubt the soundness of the empirical studies undertaken by Weinberg *et al.* And while this isn't a "knock down" conceptual point, it places the onus or burden of argumentation back upon the objectors.

(b). We can't be certain that the variations in the relevant surveys undertaken by Stich *et al.* were not influenced by extraneous factors or irrelevant details. This point is underlined when we realize that reliability of application intuition may be included by the ability of the agent to screen out irrelevant or superfluous detail: without precise control of the relevant variables (something that Weinberg *et al.* *didn't* do), there is no reason to trust data yielded from the relevant surveys.

I conclude here that, given my Mentalist account, we have good *prima facie* grounds to reject the experimental philosophy attack on intuition as a source of basic evidence. The fact of the matter, I think, is that the arguments experimental philosophers have made on this score are simply not strong or persuasive enough. And while there may be something to the project, it isn't in its present guise a threat to a priori analysis with its reliance on application intuitions, as it appears in SPM.

4.5 Conclusion

Let me now conclude this chapter. My broad aim has been to defend the claim that Moderate Naturalism is an explanatorily superior account of a priori analysis – better than either of its Incompatibilist rivals.

¹⁴³ See Jennifer Nagel. 'Epistemic Intuitions'. *Philosophy Compass*, 2 (6), 2007, pp. 792–81 and also her 'Intuitions and Experiments: A Defense of the Case Method in Philosophy'. *Philosophy and Phenomenological Research* 85 (3), 2012, pp. 495–527.

To summarize the discussion, I think that we have very good reason to believe that Mentalism does a better job of both (i) accounting for the role of a priori analysis in our SPM, which takes application intuitions as evidentially basic states, and (ii), giving us a naturalistically acceptable and metaphysically parsimonious account of concept deployment than either of its Incompatibilist rivals.

I also anticipated and provided brief replies to a set of objections in section (4.4). In order, they were:

(1) *The Subjectivity Objection*: Some may worry that Mentalism renders a priori analysis trivial or uninteresting, insofar as it makes such analysis of no significant interpersonal interest. I argue that this worry is misplaced because it overlooks the social or communal way in which even personal, psychological concepts are acquired.

(2) *The Reliability Objection*: Some experimentally-minded philosophers have also worried about the putative reliability of application intuitions. I responded in two ways: (A): that opacity of causal pathway is no more intractable than similar problems facing memorial and perceptual processes, and so we have no reason to hold the latter reliable and the former unreliable without a further argument; (B): the cultural variability objection also fails, because it is not at all clear whether the data used to support it really has the anti-SPM upshot that the experimental philosophers have claimed it does.

If I am correct here, then, Moderate Naturalism in the form of Goldman-style Mentalism has the resources both to give us a positive account of a priori analysis as it occurs in

SPM which is more satisfying than its main rivals, and to effectively respond to some of the most pressing objections one might have.

Chapter 5: Conclusion

5.1 Review of the Basics

With both the positive and negative elements of my argument for Moderate Naturalism now complete, I want to situate these elements vis-a-vis the taxonomy outlined in the first chapter, and make some concluding remarks regarding how these arguments fit into the primary aims of this thesis.

First, let us recall the basic taxonomy of the naturalism debate I constructed in the first chapter. I classified three respective metaepistemological positions vis-a-vis their stance regarding the compatibility (or incompatibility) of the following doctrines:

- (i) Naturalism (a commitment to the Metaphysical Thesis and to the Methodological Thesis).
- (ii) A Prioricity (a commitment to a priori justification).
- (iii) Conceptual Analysis (a commitment to the viability of conceptual analysis as it appears in SPM).

Incompatibilism

(*Kornblith*) Strong Naturalism – Deny (ii)-(iii), Affirm (i).

(*BonJour*) Non-Naturalism – Deny (i), Affirm (ii)-(iii)

Compatibilism

(*Goldman*) Moderate Naturalism – Affirm (i)-(iii)

My fundamental aim in introducing this taxonomy has been to motivate my defense of a Goldman-style Moderate Naturalism. More specifically, I have argued in favour of the following two-pronged thesis: (1) that both Kornblith and BonJour's attempt to establish Incompatibilism fail, and (2) that Moderate Naturalism is explanatorily superior to its Incompatibilist rivals.

Let me provide a more detailed look at each prong of my thesis, and draw out its upshot for the naturalism debate in epistemology.

5.2 *The First Prong (The Negative Project)*

The first prong of my thesis was concerned with attacking the Incompatibilist arguments promoted by Kornblith and Bonjour. In short, my aim was to undermine justification for the position that naturalism and a priori analysis are somehow mutually exclusive. Let us examine my argument.

To start, consider what Incompatibilism is claiming by considering the principle I take to exemplify the primary commitments of Incompatibilism:

(ICA): Epistemological naturalism (i) and a priori analysis (ii)-(iii) are mutually exclusive: to adopt naturalism is to reject a priori analysis and vice versa.

It is my view that we have good reason to doubt the defenses of the *(ICA)* principle we found in Kornblith and Bonjour. Neither epistemologist succeeds in the attempt to establish the mutual exclusion of (i) naturalism and (ii)-(iii) a priori analysis. Because neither of these two unlikely allies is able to make a persuasive case for Incompatibilism, I believe that the *(ICA)* principle is insufficiently motivated and we are therefore warranted in rejecting it. Let us now review my objections to each author in turn:

Against Kornblith's Exclusionary Argument: To review, Kornblith's naturalistic interpretation of the *(ICA)* principle lead him to advocate a view in which the adoption of (i) rationally requires a rejection of (ii)-(iii). Roughly, the Exclusionary Argument contends that we must reject a priori analysis as valuable because it fails to square with the correct target domain of epistemological inquiry – natural kinds – and is thus a useless method.

I advanced two main objections to Kornblith. First, it appears that the reductionism implied by his view of the correct target domain is false; or at any rate, we

have no good reason to accept it. And because it is reductionism which justifies Kornblith's methodological sanction against a priori analysis, the motivation for the Exclusionary Argument is severely undercut (for without the support of the key reductive premise Kornblith's prohibition appears totally unmotivated). Second, I argued that Kornblith's methodology is vulnerable to sceptical worries, and as such we have a general abductive case against it.

Thus, because Kornblith motivates his methodological prohibition against (ii)-(iii) via the implausible, and probably untenable assumption of reductionism. I believe that we are warranted (on pain of general scepticism) in rejecting it. Thus, I conclude, Kornblith has not succeeded in making the case for Incompatibilism.

Against BonJour's Deficiency Argument: BonJour advocates a Non-Naturalist interpretation of the (ICA) principle in which the adoption of (ii)-(iii) rationally demands a rejection of (i). His argument, in brief, was that naturalism can't justify principles of inference; thus, to avoid a radical scepticism which results from seeing all principles of inference as unjustified, we must adopt a Non-Naturalist, rationalist conception of a priori analysis.

My objection to BonJour's Deficiency Argument went as follows: we have no good reason to think that countenancing an account of a priori justification commits us to some type of Non-Naturalist Incompatibilism (viz. rationalism). This is because BonJour's supporting argument begs the question at the key juncture, and hence appears to be motivated not by any metaepistemological realities pertaining to the a priori, but rather by his own previous commitments to Non-Naturalist (Platonic) rationalism.

The bottom line, I think, is that we are warranted in rejecting Bonjour's Non-Naturalist endorsement of Incompatibilism. For his main argument simply doesn't succeed in demonstrating that naturalism implies scepticism. Thus, we are not committed to the rejection of (i) as a consequence of adopting (ii)-(iii).

In sum, I argue that my objections to Kornblith and Bonjour advanced in Chapters 2 and 3 give us good reason to reject the Incompatibilist (*ICA*) principle on grounds of being insufficiently motivated. And as a result, I believe that neither Kornblith nor Bonjour has succeeded in undercutting a Goldmanian reconciliation project.

5.3 The Second Prong (The Positive Project)

Demonstrating that we have good reason to reject the anti-Compatibilist (*ICA*) principle is, in and of itself, very important because it opens up the logical and argumentative space for a possible synthesis of (i) and (ii)-(iii). The second prong of my thesis built upon this logical space and argued that one such Compatibilist position, i.e. Moderate Naturalism, is not only possible but (*vis -à vis* the relevant doctrines) *superior* to both its Strong-Naturalist and Non-Naturalist rivals.

The case is presented through the arguments I made in Chapter 4. To review, the first argument maintained that only a Mentalist approach to philosophical analysis adopted by Moderate Naturalism can give us a plausible story as to how application intuitions, in conformity with what I have called our 'Standard Philosophical Methodology', could be an evidentially basic reliable indicator (*RI*) of the concepts which underpin their deployment. The second argument claimed that Moderate

Naturalism is additionally the more parsimonious metaepistemology, on account of its Mentalist construal of the relevant target domain of philosophical analysis.

Thus, on strength of these two arguments, I claim that Moderate Naturalism, in opposition to either Kornblith's Strong Naturalism or Bonjour's Non-Naturalism, is the preferable account of a priori concept deployment

5.4 Summarizing Remarks

To finish, let me summarize the broadest contours of my thesis. I believe that I have made a compelling case for my twin-pronged argument which, to review, is composed of the following two claims: first, that we are warranted in rejecting the Incompatibilist arguments advanced by Kornblith and Bonjour (and thus the *(ICA)* principle that both are assuming), and second, that we have good reason to think that a Compatibilist, Moderate Naturalism is the superior metaepistemological choice. In addition, I have also constructed a taxonomy of positions in the naturalism debate which, I think, should be valuable even if one isn't convinced by my case for Moderate Naturalism.

My hope is that I have made some progress with regard to clarifying the basic contours of the naturalism debate in epistemology, where optimally this clarification can contribute (however modestly) towards a more detailed future exposition and defense of Moderate Naturalism as a worthwhile *via media*. In any case, I believe that I have been successful in meeting the primary aims I set out to achieve.

Works Cited

Baker, Alan, "Simplicity", *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2011/entries/simplicity/>.

Bealer, George. A Theory of Concepts and Concept Possession'. *Philosophical Issues*, 9, 1998, pp. 261-301.

Bealer, George. 1992. 'The Incoherence of Empiricism', *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol.66, 1987, pp. 99-143.

Benacerraf, Paul. 'Mathematical Truth'. *Journal of Philosophy* 70 (19), 1973, pp. 661-679.

Bickle, John, "Multiple Realizability", *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/multiple-realizability/>.

Bickle, John. *Psychoneural Reduction*, Cambridge MA: Bradford Books/MIT Press, 1998.

Bird, Andrew. 'Necessarily, Salt Dissolves In Water'. *Analysis* 61 (4), 2001, pp. 267–274.

Block, Ned. 'Advertisements for a Semantics for Psychology'. *Midwest Studies in Philosophy* 10 (1), 1986, pp. 615-78.

Block, Ned. 'Semantics, Conceptual Role. In Edward Craig (ed.), [Book Chapter] (Unpublished). London UK: Routledge Publishing Company, 1997.

Boghossian, Paul. 'Inference and Insight'. *Philosophy and Phenomenological Research* 63 (3), 2001, pp. 633–640.

BonJour, Laurence. 'Against Naturalized Epistemology'. *Midwest Studies in Philosophy* 19, (1), 1994, pp. 283-300.

BonJour, Laurence. *In Defense of Pure Reason*. Oxford UK: Oxford University Press, 1998.

BonJour, Laurence. *The Structure of Empirical Knowledge*. Cambridge MA: Harvard University Press, 1985.

Boyd, Richard, 'How to be a Moral Realist'. In G. Sayre-McCord (ed.), *Moral Realism*. Cornell University Press, 1988.

Boyd, Richard, 'Materialism Without Reductionism: What Physicalism Doesn't Entail'. In Ned

Block (ed.), *Readings in the Philosophy of Psychology*. , Vol 1, 1980.

Brown, Curtis, "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/content-narrow/>.

Bryson, Anthony and David Alexander, 'The View from the Armchair: Responding to Kornblith's Alternative to Armchair Philosophy,' *Essays in Philosophy*, 13 (1), 2012, pp. 161-181.

Burge, Tyler. 'Individualism and the Mental'. *Midwest Studies in Philosophy* 4 (1), 1979, pp. 73-122.

Brueckner, Anthony. 'Review: Laurence Bonjour, In Defense Of Pure Reason'. *Nous*, 34 (2), 2000, pp. 302-311.

Casullo, Albert. *A Priori Justification*. New York: Oxford University Press, 2003.

Cheyne, Colin. *Knowledge, Cause and Abstract Objects: Causal Objections to Platonism*. Boston, MA: Kluwer Academic Publishers, 2001.

Devitt, Michael. *Designation*. New York: Columbia University Press, 1981.

Fodor, Jerry. 'Special Sciences, or the Disunity of Science as a Working Hypothesis', *Synthese*, 28, 1974, pp. 97-115.

Fodor, Jerry. *The Language of Thought*, Cambridge MA: Bradford Books/MIT Press, 1975.

Fodor, Jerry. 'The Special Sciences: Still Autonomous After All These Years'. *Philosophical Perspectives* (11), 1997, pp. 149-63

Goldman, Alvin. 'A Priori Warrant and Naturalistic Epistemology'. *Noûs*, Vol. 33, Supplement: Philosophical Perspectives, 13, Epistemology, 1999, pp. 1-28.

Goldman, Alvin. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press, 1986.

Goldman, Alvin. 'Epistemic Folkways and Scientific Epistemology'. In *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. Cambridge MA: MIT Press, 1992, pp.155-175

Goldman, Alvin. 'Philosophical Intuitions: their target, their source, and their epistemic status'. *Grazer Philosophische Studien*, 74, 2007, pp. 1-26.

Goldman, Alvin. 'Philosophical Naturalism and Intuitional Methodology'. *Proceedings and Addresses of the American Philosophical Association*, 2010.

Goldman, Alvin & Joel Pust. 'Philosophical Theory and Intuitional Evidence'. In M. DePaul & W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield, 1998, pp. 179-197.

Goldman, Alvin. 'What is Justified Belief?' In G.S. Pappas, ed., *Justification and Knowledge*. Dordrecht: D. Reidel Publishing Company, 1979, pp. 1-23.

Goodman, Nelson. *Fact, Fiction and Forecast*. 2nd Edition, Bradford Books/MTI Press, 1983.

Grice, H. P. *Studies in the Ways of Words*. Cambridge MA: Harvard University Press, 1989.

Horgan, Terence. 'Nonreductive Materialism and the Explanatory Autonomy of Psychology'. In Steven J. Wagner & Richard Warner (eds.), *Naturalism: A Critical Appraisal*. New York: University of Notre Dame Press, 1993.

Kim, Jaegwon. 'Multiple Realization and the Metaphysics of Reduction'. *Philosophy and Phenomenological Research*, 52 (1), March 1992, pp. 1-26.

Kornblith, Hilary. *Knowledge and its Place in Nature*. Oxford UK: Clarendon Press, 2002.

Kornblith, Hilary. 'In Defence of a Naturalized Epistemology'. In Ernest Sosa and John Greco (Eds.), *The Blackwell Guide to Epistemology*, Blackwell Publishing: Malden MA, 1998, pp. 158-169.

Kornblith, Hilary. 'Naturalism, Both Metaphysical and Epistemological.' *Midwest Studies in Philosophy* 19 (1), 1994, pp. 39-52.

Kornblith, Hilary, Ed. *Naturalizing Epistemology*. Cambridge MA: MIT Press, 1985

Kornblith, Hilary. 'Naturalism and Intuitions.' In C. Beyer & A. Burri (eds.), *Philosophical Knowledge: Its Possibility and Scope*. Grazer Philosophische Studien, Amsterdam: Rodopi, 2007, pp. 27-49.

Kornblith, Hilary. 'Replies to Alvin Goldman, Martin Kusch and William Talbott.' *Philosophy and Phenomenological Research*, 52, (2), 2005, pp. 427-441.

Kripke, Saul. *Naming and Necessity*. Harvard MA: Harvard University Press, 1980.

Lewis, David. *On The Plurality of Worlds*. Princeton NJ: Princeton University Press, 1986.

Liggins, David. 'Epistemological Objections to Platonism.' *Philosophy Compass*, 5, (10), 2010, pp. 67-77.

Marr, David. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York NY: Freeman Books, 1982.

McGinn, Colin. 'A Note on the Essences of Natural Kinds.' *Analysis*, (35), 6, June 1975, pp. 178-180.

Nagel, Ernest. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York NY: Harcourt, Brace & World, 1961.

Nagel, Jennifer. 'Epistemic Intuitions.' *Philosophy Compass*, 2 (6), 2007, pp. 792-81.

Nagel, Jennifer. 'Intuitions and Experiments: A Defense of the Case Method in Philosophy.' *Philosophy and Phenomenological Research* 85 (3), 2012, pp. 495-527.

Papineau, David, "Naturalism", *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/entries/naturalism/>.

Papineau, David. *Philosophical Naturalism*. London, UK: Blackwell Publishing Co, 1993.

Peacocke, Christopher. *A Study of Concepts*, Cambridge MA: Bradford Books/MIT Press, 1992.

Peacocke, Christopher. 'Three Principles of Rationalism'. *European Journal of Philosophy*, (10), 3, 2002, pp. 375-397.

Plantinga, Alvin. *Warrant, the Current Debate*. New York NY: Oxford University Press, 1993.

Putnam, Hilary. 'Psychological Predicates', in W. H. Capitan and D. D. Merrill, eds., *Art, Mind and Religion*, Pittsburgh PA: University of Pittsburgh Press, 1967.

Putnam, Hilary. *Realism and Reason: Philosophical Papers, Vol. 3*. Cambridge UK: Cambridge University Press, 1983.

Putnam, Hilary. 'The Meaning of 'Meaning''. *Minnesota Studies in the Philosophy of Science*, (7), 1975, pp. 131-193.

Pylyshyn, Zenon. *'Computation and Cognition'*, Cambridge MA: Bradford Books/MIT Press, 1984.

Quine, W.V.O. 'Epistemology Naturalized'. In *Ontological Relativity and Other Essays*, New York: Columbia University Press, 1969.

Rey, Georges. 'Digging Deeper for the A Priori'. *Philosophy and Phenomenological Research*, 43 (3), November, 2001, pp. 649-656.

Russell, Bertrand. *The Problems of Philosophy*. Oxford UK: Oxford University Press, 1912.

Sober, Eliot and David Sloan Wilson. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard MA; Harvard University Press, 1998.

Tye, Michael. 'Naturalism and the Mental'. *Mind* 101 (403), 1992, pp. 421-441.

Weinberg, Jonathan M., Nichols, Shaun M. and Stich, Stephen. 'Normativity and Epistemic Intuitions'. *Philosophical Topics*, 29 (1-2), 2001, pp. 429-460.

Weller, Cass. 'BonJour and Mentalese'. *Synthese* 113 (2), 1997, pp. 251-63.