Challenging Exclusionary Naturalism

Nathan Robert Cockram

Department of Philosophy, Carleton University

The purpose of this paper is to reconstruct Hilary Kornblith's (2002) argument for excluding conceptual analysis from epistemological inquiry, and then provide three objections to it. More specifically, Kornblith argues that epistemological properties such as *knowledge* reduce to natural kinds (with a constitutive essence) which can only be discovered and investigated using the a posteriori methods of the natural sciences. Thus, he continues, conceptual analysis cannot properly illuminate the target domain. The three objections to Kornblith's argument which I present are as follows: (i) Multiple Realizeability, (ii) Psychological Explanation, (iii) Starting Points. On strength of these objections, I conclude that Kornblith's brand of a posteriori epistemology both eliminates our ability to make epistemic evaluations in general, and also implies a strong form of scepticism.

Keywords: Kornblith, conceptual analysis, naturalized epistemology

1. Introduction

Hilary Kornblith (2002) argues in favour of a metaepistemological thesis which I call 'Exclusionary Naturalism' (EN). EN contends that the relevant target domain of philosophical analysis is populated by empirical natural kinds. Kornblith takes this thesis as placing significant adequacy constraints upon any naturalistically respectable metaepistemology for the following reason. It is generally accepted that consulting our a priori application intuitions regarding a predicate can only generate and justify inferences ranging over what is conceptually possible. According to Kornblith, however, what we are *really* interested is an a posteriori and empirical account of the phenomenon as it exists here "in the world" (Kornblith 2002, 13–20). Thus—and this is the crux of EN—Kornblith denies that the deliverances of a priori analysis can be justified vis-à-vis the relevant target domain.[1] While we will

Corresponding author's address: Nathan Robert Cockram, Carleton University, 1125 Colonel By Dr, Ottawa, ON K1S 5B6, Canada. Email: nrcockerham86@gmail.com.

[1] Note that Kornblith does not 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

Studia Philosophica Estonica (2014) **7.1**, 1–34
Published online: March 2014
Online ISSN: 1736–5899
www.spe.ut.ee
http://dx.doi.org/10.12697/spe.2014.7.1.01

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. My primary aim in this paper will be to show that the general line of argumentation Kornblith recruits in favour of the EN thesis—which I will call the 'Exclusionary Argument' (EA), fails to do the justificatory work required in order to establish EN. Thus, the conclusion I wish to establish is that his a posteriori program in epistemology poses no significant threat to the use of application intuitions or intuitional methodology.

My argument will proceed in two stages: first, (section 2) I will examine Kornblith's Exclusionary Argument as presented in his 2002 book. Second, I will then proceed (sections 3-5) with arguments designed to show that it does not successfully establish what Kornblith needs in order to set his Exclusionary Naturalism on firm ground.

2. The Exclusionary Argument Reconstructed

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 (scientific)[2] methodological purview and must be discarded.

To begin, consider the following argument, which I believe is a very plausible reconstruction of EA:

(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.

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).

[2] For more, see (Papineau 2009).

(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 alone cannot illuminate the natural kind(s) that stand as the referential content of the term(s) contained within its target domain.[3]

(6) Therefore, a priori analysis is a scientifically suspect method.

(7) Therefore, epistemology cannot retain a priori analysis.

Reconstructing EA 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 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.

2.1 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.

[3] Thanks to an anonymous referee for suggesting that I add the qualifier 'alone' to this premise. This, I think, both renders my argument more faithful to Kornblith's position, and at the same time makes it more substantive.

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 plays an integral and a unifying role in the prediction and explanation of the successful adaptation of organisms to environmental requirements.

2.1.1 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.[4] Given that many organisms are successful in meeting these demands, what would adequately explain this capacity?

Kornblith (1985, 4f) 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.

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 would not). 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 (Kornblith 2002, 56–61).

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 beliefs about their environment—producing true beliefs is conducive to survival and reproduction (Kornblith 2002, 56–59).

[4] 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).

2.1.2 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*…(Kornblith 2002, 57–58, italics added)

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-à-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 type* (Kornblith 2002, 61–62), 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.[5]

2.1.3 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.

[5] See Kornblith 2002, 61–69

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";[6] 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 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 in the prediction of behavior successful in meeting biological needs and thereby implicated in the Darwinian explanation of the selective retention of traits. (Kornblith 2002, 62)

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. (Kornblith 2002, 11, italics added)

In other words, the process-type "Knowledge" denotes an external (in the epistemological sense outlined above) set of co-occurring and mutually

[6] 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 (1980, 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 (Kornblith 2002, 61–62).

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.

2.2 The Purported Methodological Upshot

With the three-part subargument for knowledge as a natural kind outlined in 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 *cannot* 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. (Kornblith 2007, 36)

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 2.1 (c), i.e.:

(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 alone cannot illuminate the natural kind(s) that stands as the referential content of the term(s) contained within its target domain.

This cannot 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 analyzing the relevant descriptive content of the scenario in question; blocking Kornblith's desired conclusion.[7]

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 out folk chemical notions are to chemists? (Kornblith 2002, 18–19)[8]

[7] For more, see (Kripke 1980).

[8] Also, consider the following suggestive passage from (Kornblith 2005, 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

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").[9] 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 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 (H_2O or XYZ). This is why the Causal theory is essentialist: all and only those worlds containing H_2O are worlds in which 'water' refers to water, for water just *is* H_2O—'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 causal theory of reference (and by entailment property 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

properly identifies it as H_2O. 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 (Kornblith 2005).

[9] More on this below but see (Kripke 1980), (Putnam 1975, 1983).

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.[10]

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

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 EA and reject its purported methodological upshot.

My objection to the Natural Kinds account of knowledge will employ two broad lines of argumentation. First (section 4), I will grant Kornblith's position and maintain that once we tease out its strongly reductionist consequences, the Natural Kinds account is revealed to be a highly implausible

[10] More accurately, Kornblith denies that a priori conceptual analyses can be justified on the basis of application intuitions. So, he does not 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.

theory of knowledge (on grounds of explanatory adequacy). Second (section 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 than his supposed competitors.

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 (4.1), I will put forth two anti-reductionist arguments adopted from similar debates in the philosophy of mind.[11] First (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 arbitrary and most likely false. Second (4.3), I argue that reductionism about knowledge has the further undesirable consequence in that it arbitrarily stops us from making important theoretical generalizations about the epistemic property denoted by 'Knowledge'.

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.[12]

A good way to flesh out the relationship is to consider the modal status of primary and secondary nomic properties.[13] Contra descriptivism, Kripke

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

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

[13] 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

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 $P_1, \ldots, P_n$ across all possible worlds.[14]

(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).[15]

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.[16] Thus, reductionism about 'knowledge' can be defined as such:

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 (McGinn 1975).

[14] For a similar formulation, see (Bird 2001).

[15] 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 H_2O). Thus, Water and H_2O are identical (necessarily coextensive) and related rigidly: there is no possible world in which Water is not H_2O and conversely, there is no possible world in which H_2O lacks the secondary properties characteristic of water. See (McGinn 1975, 181). Also, consider the following passage from Kripke's (1980, 125): "In particular, then, present scientific theory is such that it is 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 unquestionably are not part of the *meaning* of 'gold' and were not known with *a priori* certainty)."

[16] Intension and extension collapse in the case of necessary coextensions because in such

(RK) The natural kind-term 'knowledge' is a (directly denoting) semantic primitive which is necessarily coextensive with the same set of essential physical properties $P_1, \ldots, P_n$ 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.

4.2 The Multiple Realizabilty 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 vocabulary,[17] are necessarily coextensional with lower-order physical predicates. And it is this semantic-metaphysical move, I emphasize, which provides the foundation for Kornblith's methodological sanction on the a priori.[18] As a result, if it could be determined that the kind identity in question *is false*,

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

[17] 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 1992), (Block 1997).

[18] 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 the soundness of his natural kinds argument for property essentialism about knowledge.

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.[19]

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 *cannot* be identical with *P*. So 'knowledge' *is not* 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.

(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.[20]

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 Putnam (1967), and the conceptual variant is derived from an argument made by Jerry Fodor (1974, 97–115).

4.2.1 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

[19] 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 (Kim 1992).

[20] For a similar argument, see (Bickle 2013). Also see (Horgan 1993).

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 molluscs, and certainly feel pain), etc. (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 realizing physiological base properties differ. And according to Putnam, (given neurobiological structural diversity) it is "overwhelmingly likely" (Putnam 1967, 437)[21] 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*.

4.2.2 Fodor's Argument

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.

[21] See also (Bickle 1998, 116–117).

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 every event which falls under a law of psychology *also* falls under a law of physics:[22]

> 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. (Fodor 1974, 132–133)

The problem, argues Fodor, is that the possibility of multiple realization[23] *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) $(K \leftrightarrow (P_1 \vee P_2 \vee P_3 \vee P_4 \vee \ldots \vee P_n))$

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.[24] The idea here, I think, is that we

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

[23] Here Fodor (1974) cites neuroplasticity data taken from the work of Karl Lashley. Also see (Fodor 1975).

[24] That is, projectability, according to Kim (1992, 11): "…is the ability to be confirmed by observation of positive instances such that any generalized conditional of the for "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

cannot 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).[25] 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-à-vis their lower level physiological correlates; an autonomy which establishes the falsity of reductionism.[26]

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) kinds,[27] 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

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."

[25] Here is how Kim (1992, 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 1983).

[26] Thus Fodor (1974, 13): "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.

[27] I take it that this is not a controversial point: what, other than token intentional states, could epistemic states be?

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

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 (1984). 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.

4.3.1 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:

> 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. (Bickle 1998, 152)

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) $(\forall S)$ $(\forall D)$ $(\forall A)$ (If S wants D and believes that doing A will bring about D, then ceteris paribus S will do A).[28]

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 know 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 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

[28] See (Horgan 1993, 298).

> 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. (Pylyshyn 1984, 11)

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).

As an aside, note that the argument from Psychological Explanation, although invoking the premise of multiple base realizers is, I believe, 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).

4.3.2 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 it 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.

4.3.3 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':[29]

(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.[30] 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.[31]

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 do not support projectable counterfactuals. This is a significant problem for Kornblith's natural kinds account. For if the reductive view of epistemic properties cannot capture higher level generalizations such as (EG), then it lacks explanatory power vis-à-vis its initial target *explanandum* and thus appears to be self-defeating.

4.3.4 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.

[29] For more on epistemic states as propositional attitudes, see (Horgan 1993).

[30] 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.

[31] See (Bickle 1998, 152) and (Kim 1992), (Horgan 1993).

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).[32]

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.[33] And it is adopting a causal theory of reference which justifies 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.[34] Thus, modifying the relevant premise of the Natural Kinds subargument in the face of the projectability problem does not seem like a live option for Kornblith, lest he was to abandon his overall metaepistemological aims.

[32] See footnotes 12–14 above.

[33] 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 identical to* a set of primary properties *P* which constitute or "give rise to" *S*. For more details, see footnote 14 above and (Putnam 1975).

[34] 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).

4.3.5 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 cannot be 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).[35]

[35] Thanks to David Matheson for suggestions here.

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.

5. The "Starting Points" Objection[36]

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.

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-à-vis epistemological kinds and the empirical method he puts in its place. The root of Kornblith's rejection follows from his view of epistemological inquiry as directed at extramental natural kinds. More specifically, he *denies that conceptual analysis can be justified on the basis of application intuition* (Kornblith 2002, 10–11).

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. (Kornblith 2002, 10–11)

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

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 ostensivley "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*.[37]

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.

5.2 The Disanalogy Premise and Knowledge

Kornblith claims that we can ostensively "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).

5.2.1 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?[38] Thus, on further examination Kornblith's analogy to the rock collector highlighted above breaks down into a *disanalogy*—for even if we ignore the considerable issue of theory-ladenness in observability

[37] This is not a commitment rendered explicit by Kornblith, but it is implicit within his discussion. See (Kornblith 2002, chapters 1–2).

[38] Note that this problem of identification is orthogonal to what Devitt calls the "Qua Problem" for causal theories of reference. See (Devitt 1981).

(which the Initial Designation stage seems to ignore),[39] it does not seem as if we could ostensivley "point to" knowledge:

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

The reason for the disanalogy seems clear: unlike water or gold, knowledge lacks an empirical "mode of presentation" through which we could associate a kind term with an extensional reference class via an initial baptismal act. In other words, we cannot just look into the world, point and say "That *such-and-such x* shall be designated with the kind term 'Knowledge'".

5.2.2 Fixing the Reference of Knowledge

Assuming that the Disanalogy Premise is correct, Kornblith cannot simply "point to" and ostensivley 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).[40]

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

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,[41] 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

[39] As an anonymous referee correctly pointed out, Kornblith's ES implicitly assumes a naive empiricism about observable properties. And while I agree that theory-ladenness causes considerable problems for this assumption, I think the argument is stronger as it stands, without relying on considerations of theory-ladenness. For we can see the Disanalogy Premise as claiming that, even if we ignore the substantial problems theory-ladeness poses to his natural kinds view (i.e. even if we grant naive empiricism), Kornblith's anthology *still* breaks down because 'Knowledge' does not have any observational properties that could even be theory-laden to begin with.

[40] See (Kornblith 2002, 63–69).

[41] 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.

knowledge (vis-à-vis a set of hypothetical scenarios).[42] Thus, despite his explicit endorsement of ES, Kornblith cannot in practice 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.

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 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 do not 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 judgements. (Bryson and Alexander 2012, 15)

Because we can identify the latter kinds via observational properties, zeroing in on the correct explanandum is both relatively straightforward and uncontroversial. But this is not at all the case with knowledge: we must rely upon

[42] 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.

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 reliabilst 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. Thus, Kornblith *is not* 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 does not seem that he—or anyone else for that matter—will be able to successfully illuminate the target phenomenon. Summarized, The Starting Points Objection is this:

(1) As per the DP, 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-à-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 or tracks an extramental natural kind (or has any relevance to an extramental domain of inquiry).

If this argument is sound, Kornblith's actual method (as opposed to his official one) is open to the very same objections he makes against those who explicitly advocate an armchair approach. For despite the emphasis Kornblith places upon ES, he advocates an analysis of knowledge which, rather than being causal/ostensive, proceeds a priori. As a result, Kornblith's methodological sanction against the a priori serves to defeat *his own* reliabilist theory in the same manner that it purports to defeat his armchair rivals; for if knowledge is in fact a natural kind, then we have no reason to think that his armchair reliabilism picks out the correct properties a priori. In fact, I think that the argument has an even stronger upshot, which is as follows: if the Disanalogy Premise is true, and we cannot pick out instances of knowledge by empirical means, then it follows that *any* proposed analysis

of knowledge will be defeated (i.e. ruled out as unjustified) by ES, because the latter methodology permits all and only those analyses which make empirical (i.e. causal/ostensive) identities.

Thus, for the reasons given above, if the Starting Points argument is on the mark, Kornblith's Empirical Strategy is tantamount to a form of scepticism regarding epistemic property ascriptions. Moreover, given that the adoption of such a sceptical metaepistemology would have highly implausible consequences for the practise of normative epistemology (which unfortunately I do not have space to elaborate upon here),[43] I believe that we therefore have good reason to reject the ES outright.

6. Conclusion

As a means to concluding this paper, 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.

6.1 Summary of Kornblith's Argument

In the broadest terms, Kornblith's primary aim 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 *is not* amenable to conceptual analysis.

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

(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.

[43] For more on this, see (Cuneo 2007) and (Lynch 2009).

(5) *A priori* analysis alone 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 cannot illuminate the target explanandum.

6.2 Objections and 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 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 a lower level reducing (physical) kind. 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.

In sum, I think that these three objections give us good reason to reject Kornblith's Exclusionary Argument. That is, once we draw out the reductionist tendencies of his a posteriori brand of epistemology, it appears to be highly implausible, if not altogether untenable.

Acknowledgements

I would like to thank David Matheson and Gordon Davis for comments.

Bibliography

Bealer, G. (1996). *A priori* knowledge and the scope of philosophy, *Philosophical Studies* **81**: 121–142.

Bickle, J. (1998). *Psychoneural Reduction*, Bradford Books/MIT Press, Cambridge, MA.

Bickle, J. (2013). Multiple realizability, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2013 edn.
URL: http://plato.stanford.edu/archives/spr2013/entries/multiple-realizability/

Bird, A. (2001). Necessarily, salt dissolves in water, *Analysis* **61**: 267–274.

Block, N. (1997). *Semantics, Conceptual Role*. unpublished book chapter.

BonJour, L. (1998). *In Defense of Pure Reason*, Cambridge University Press, Cambridge.

Boyd, R. (1980). Materialism without reductionism: What physicalism does not entail, *in* N. Block (ed.), *Readings in Philosophy of Psychology*, Harvad University Press, Cambridge, MA, pp. 67–106.

Boyd, R. (1988). How to be a moral realist, *in* G. Sayre-McCord (ed.), *Moral Realism*, Cornell University Press, Ithaca, pp. 181–228.

Bryson, A. and Alexander, D. (2012). The view from the armchair: Responding to Kornblith's alternative to armchair philosophy, *Essays in Philosophy* **13**: 161–181.

Burge, T. (1979). Individualism and the mental, *Midwest Journal of Philosophy* **4**: 73–122.

Cuneo, T. (2007). *The Normative Web*, Oxford University Press, New York.

Devitt, M. (1981). *Designation*, Columbia University Press, New York.

Fodor, J. (1974). Special sciences, or the disunity of science as a working hypothesis, *Synthese* **28**: 97–115.

Fodor, J. (1975). *Language of Thought*, Harvard University Press, Cambridge, MA.

Goodman, N. (1983). *Fact, Fiction and Forecast*, 2nd edn, Bradford Books/MIT Press, Cambridge, MA.

Horgan, T. (1993). Nonreductive materialism and the explanatory autonomy of psychology, *in* S. J. Wagner and R. Warner (eds), *Naturalism: A Critical Appraisal*, University of Notre Dame Press, New York, pp. 295–320.

Kim, J. (1992). Multiple realization and the metaphysics of reduction, *Philosophy and Phenomenological Research* **52**: 1–26.

Kornblith, H. (2002). *Knowledge and its Place in Nature*, Clarendon Press, Oxford.

Kornblith, H. (2005). Replies to Alvin Goldman, Martin Kusch and William Talbott, *Philosophy and Phenomenological Research* **52**: 427–441.

Kornblith, H. (2007). Naturalism and intuitions, *in* C. Beyer and A. Burri (eds), *Philosophical Knowledge: Its Possibility and Scope*, Rodopi, Amsterdam, pp. 27–49.

Kornblith, H. (ed.) (1985). *Naturalizing Epistemology*, MIT Press, Cambridge, MA.

Kripke, S. (1980). *Naming and Necessity*, Harvard University Press, Cambridge, MA.

Lynch, M. (2009). Truth value and epistemic expressivism, *Philosophy and Phenomenological Research* **79**: 76–97.

McGinn, C. (1975). A note on the essences of natural kinds, *Analysis* **35**: 178–180.

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

Papineau, D. (2009). Naturalism, *in* E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
URL: http://plato.stanford.edu/archives/spr2009/entries/naturalism/

Putnam, H. (1967). Psychological predicates, *in* W. H. Capitan and D. D. Merrill (eds), *Art, Mind and Religion*, University of Pittsburgh Press, Pittsburgh, pp. 37–48.

Putnam, H. (1975). The meaning of 'meaning', *Minnesota Studies in the Philosophy of Science* **7**: 131–193.

Putnam, H. (1983). *Realism and Reason: Philosophical Papers, Vol. 3*, Cambridge University Press, Cambridge.

Pylyshyn, Z. (1984). *Computation and Cognition*, Bradford Books/MIT, Cambridge.

Quine, W. (1969). Epistemology naturalized, *Ontological Relativity and Other Essays*, Columbia University Press, New York, pp. 69–90.

Sober, E. and Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behaviour*, Harvard University Press, Cambridge, MA.

Tye, M. (1992). Naturalism and the mental, *Mind* **101**: 421–441.