Is Understanding Explanatory or Objectual?
[Forthcoming in *Synthese*. Please cite final draft.]

Kareem Khalifa
Department of Philosophy
Middlebury College
Middlebury, VT 05753

Abstract: Jonathan Kvanvig has argued that "objectual" understanding, i.e. the understanding we have of a large body of information, cannot be reduced to explanatory concepts. In this paper, I show that Kvanvig fails to establish this point, and then propose a framework for reducing objectual understanding to explanatory understanding.

Introduction

Many feats of understanding strike us as significant cognitive achievements. To choose but a few examples: Einstein's understanding of physics, Darwin's understanding of why species are adapted to their environments, a detective's understanding of how a crime was committed, and our friendly neighborhood mechanic's understanding of our car's engine.

Looking at these kinds of examples, it's easy to see why recent work in both epistemology and philosophy of science has flagged understanding as worthy of philosophical study[1]. Despite this newfound enthusiasm, understanding remains only faintly understood. Unlike neighboring concepts such as knowledge and explanation, nothing approaching an analysis of understanding has even been suggested by its recent friends. This has not deterred key players in this burgeoning "philosophy of understanding" from making bold statements. Notably, Jonathan Kvanvig—arguably the foremost

[1] Representative works in epistemology include (Elgin 2006, 2007, 2009; Kvanvig 2003, 2009a, 2009b; Pritchard 2008, 2009, 2010; Riggs 2003, 2009; Zagzebski 2001); in philosophy of science, (de Regt 2004, 2009a; de Regt and Dieks 2005; de Regt, Leonelli, and Eigner 2009; Grimm 2006, 2008, forthcoming; Lacey 1999; Khalifa forthcoming; Trout 2002, 2005, 2007). I will not include the linguistic understanding discussed by philosophers of language, e.g. (Gross 2005; Longworth 2008; Pettit 2002) in my paper.

epistemologist working on understanding—has claimed that a form of understanding that he calls "objectual understanding" cannot be reduced to another form of understanding that he calls "explanatory understanding[2]."

In this paper, I will argue that Kvanvig has failed to establish this antireductionism[3]. To that end, after reviewing Kvanvig's position (§1), I will argue that Kvanvig's antireductive arguments are inadequate (§2). I will then propose a new reductionist program (§3). If my claims are correct, I hope to have made some progress towards a philosophical analysis of understanding.

1. Kvanvig on understanding

To begin, I discuss Kvanvig's distinctions between objectual and explanatory understanding (§1.1), and then present his challenge to those who would reduce objectual understanding to its explanatory counterpart (§1.2).

1.1. Objectual and explanatory understanding

For Kvanvig, understanding has many characteristics that apply to both its objectual and explanatory flavors. Identifying these similarities will help to isolate the relevant differences. Specifically, Kvanvig (2009a, 97) takes both forms of understanding to involve a "grasp of the structural relationships (e.g. logical, probabilistic, and explanatory relationships) between the central items of

[2] (Kvanvig 2009a, 100-102, 108-109)
[3] Hereafter, all references to "reductionism" and its cognates refer to the position that objectual understanding reduces to explanatory understanding.

information regarding which the question of understanding arises[4]."
Additionally, he takes understanding of all sorts to be unaffected by epistemic luck (Kvanvig 2003, 196-200; 2009a, 103-109) and to trade in mostly true information (Kvanvig 2003, 190-191; 2009a, 97; 2009b, 341-343).

Given these commonalities, we can safely put aside further details about the nature of "grasping," the factivity of understanding, the semantics and metaphysics lurking behind Kvanvig's account of information, and understanding's immunity to Gettier luck when contrasting explanatory and objectual understanding. Instead, we will focus on their central difference—the kinds of "structural relationships" required by explanatory understanding are more restrictive than those required by objectual understanding.

Unsurprisingly, explanatory understanding requires some of these relationships to be explanatory. Canonical attributions of explanatory understanding reflect this point. Such attributions are of the form "*S* understands *why p*," where *S* refers to a person and *why p*, to an embedded question[5]. Since explanations are frequently identified with answers to why-questions, explanatory understanding entails that:

(1) If *S* (explanatorily) understands *why p*, there is some item of information *q* such that *S* grasps that *q explains p*.

For instance, if Adriana understands why Argentina suffered an economic crisis from 2001-2002, then presumably she possesses an explanation of this crisis, e.g.

[4] Elsewhere, Kvanvig (2003, 192) describes these explanatory, logical, and probabilistic relationships as "coherence-making." I will alternate freely between "structural" and "coherence" talk.
[5] Kvanvig (2009a, 100-101) also equates explanatory understanding with understanding-why.

that it resulted from the Argentine Currency Board's pegging the Argentine peso to the US dollar.

By contrast, objectual understanding "incorporates explanatory relations (*when they exist*)" (Kvanvig 2003, 101: emphasis added). However, when explanatory relations do not exist, objectual understanding can still be achieved by grasping other structural relationships—i.e., logical and probabilistic relationships. By contrast, explanatory understanding is ruled out *tout court* in such cases. Thus, objectual understanding is broader than explanatory understanding.

Objectual understanding's greater breadth is reflected in its grammatical differences with explanatory understanding, e.g.

> objectual understanding['s…] grammatical form takes an object, as when we say that Bas understands quantum theory or Cheney's buddies understand the cost of hunting with him (Kvanvig 2009a, 96).

Since all embedded why-questions are grammatical objects (namely certain interrogative content clauses), but not vice versa, it is natural to interpret explanatory understanding as narrower than objectual understanding. More precisely,

(2) If *S* (objectually) understands object *O*, then there are items of information *p* and *q* about *O* such that:

(a) *S* grasps that *q explains p*; or

(b) There is nothing that explains *p*, and *S* grasps that *q entails p, p entails q, q is probabilistically relevant to p*, or *p is probabilistically relevant to q*.[6]

As an example of (2.a), consider another Argentinean economist, Leticia, who objectually understands the economic crisis of 2001-2002. Since there is an explanation of the crisis, she, like Adriana, grasps the Currency Board's effect on the crisis. As such, there is not yet a reason to think that objectual and explanatory understanding are distinct.

1.2. Kvanvig's Example

However, the differences between our two modes of understanding become apparent with examples in which no explanation is to be found, i.e. examples of (2.b) in our account of objectual understanding. In Kvanvig's specific example, he invites us to consider an indeterministic system involving, *inter alia*, an electron's trajectory. Kvanvig asserts that while we cannot provide a causal explanation of why the electron went, e.g. to the left rather than the right, we can still have objectual understanding of the system:

> Where *S* is some indeterministic system, we can have objectual understanding of the system even though we cannot interpret this understanding in terms of being able to understand why things happen as they do in *S*… So objectual understanding cannot be reduced to

[6] I will leave the sense in which items of information are "about" objects unanalyzed, though I trust my uses hereafter are not contentious.

propositional understanding via appeal to 'wh-'complement attributions of understanding or explanations (Kvanvig 2009a, 101-102).

Given the account of objectual understanding from above, this implies that someone grasps a logical or probabilistic relationship linking the information that *the electron went left rather than right* to some other information about the system. Since such a relationship is presumed to be non-explanatory, it would appear that we (objectually) understand the system, but we do not (explanatorily) understand *why* the electron went left rather than right. Call this *Kvanvig's Example*. More precisely, Kvanvig's Example involves the following claims:

K1. If, for a given fact *p* about a system *S*, there is no explanation of why *p*, but *S* is nevertheless objectually understood, then objectual understanding does not reduce to explanatory understanding.

K2. There is no explanation for why the electron went left rather than right.

K3. The electron's going left rather than right is a fact about an objectually understood system.

As should be clear, K1-K3 entail that objectual understanding does not reduce to explanatory understanding.

2. Critique of Kvanvig's Example

However, Kvanvig's Example is unsound, as both K1 and K2 are highly contentious. The antecedent of K1 is compatible with even the most thoroughgoing forms of reduction. As a result, it suffers from what I call the *Failed Criterion Objection* (§2.1). Against K2, philosophical reflections on

indeterministic explanations as they are found in scientific practice strongly suggest that the probabilistic and logical relations involved in Kvanvig's Example may actually be explanatory. Call this the *Hidden Explanation Objection* (§2.2). Obviously, the point here is not so much to squabble with Kvanvig about electrons, but to highlight how the richness of the philosophical literature on explanation poses difficulties in finding unexplainable facts that are still understandable (§2.3).

2.1. The Failed Criterion Objection

For the sake of argument, let us suppose a particularly strong form of reductionism—that all understanding is explanatory. The resulting reinterpretation of Kvanvig's Example would then suggest the following:

K2*. The electron's going left rather than right is not explained, and hence not understood.

K3*. The electron's going left rather than right is a fact about an explanatorily understood system.

This appears as plausible an interpretation of Kvanvig's Example as the one suggested by K2 and K3. Given that we assumed a particularly strong form of reductionism, this provides compelling evidence that K1 is an insufficient criterion for gauging the reducibility of objectual understanding to its explanatory counterpart. That is the Failed Criterion Objection.

Admittedly, Kvanvig has not told us exactly which probabilistic and logical relationships tether the information about the electron's trajectory to the system, and moreover has not told us the precise manner in which the system is

objectually understood. However, that only reinforces the current objection, for presumably those details would force a more demanding replacement for K1.

Let me flag a potential lacuna in this interpretation. K3* refers to an "explanatorily understood system." By itself, a system is not an embedded why-question, so this expression may betray some sort of category mistake. However, I suggest that this is just shorthand for understanding why the system functions (operates, behaves, etc.) in the ways that it does, i.e. why the system "does what it does" or "is what it is."

Importantly, the Failed Criterion Objection does not entail that reductionism is correct. Rather, it simply shows that Kvanvig's criterion, K1, cannot settle that matter in one way or the other. I will discuss my own criteria for assessing reducibility in §3.

2.2. The Hidden Explanation Objection

As already noted, Kvanvig does not specify the logical and probabilistic relationships connecting the information about the electron's going left rather than right to other information about the system. However, some logical and probabilistic relationships are explanatory. As a result, antireductionists must eliminate this possibility. Failure to do so licenses the Hidden Explanation Objection. Kvanvig's Example appears susceptible to this Objection. If this is correct, then not only K1, but also K2, is false.

In critiquing K2, my emphasis is less about whether we can actually explain why the electron went left rather than right—that issue is best left to philosophers of physics. My primary aim is to highlight the contentiousness of

Kvanvig's reasons for taking a broad class of facts—which includes the fact about the electron—to be unexplainable. To that end, Kvanvig takes the conjunction of four general features to preclude the possibility of explaining why the electron went left rather than right (Kvanvig 2009a, 101-102):

(1) The explanation required is *causal*: "if there is no cause of the electron going left rather than right, there is no explanation why the electron went to the left either."

(2) The explanandum is *indeterministic*: "In indeterministic systems, things happen that are uncaused, both probabilistically and deterministically."

(3) The explanandum is *contrastive*: "the events in question are irreducibly indeterministic in such a way that there is no causal explanation as to why the actual events occurred *rather than some other events*" (emphasis added).

(4) The explanandum contrasts *equally probable* outcomes: "If the probability of an electron going to the left is precisely the same as that of going to the right (and there is no hidden variable to account for the difference), then whichever way it goes is the result of chance rather than causation."

Thus, Kvanvig is denying the possibility of causal, indeterministic explanations of explananda contrasting equally probable outcomes. However, some of these requirements are illicit, and the rest can be satisfied. I will build my case against this position incrementally, addressing each requirement in turn.

2.2.1. Causal explanation

First, observe that K2 does not require the explanation to be causal. Any acceptable kind of explanation (causal or otherwise) of the electron's trajectory is an example of explanatory understanding. Moreover, there are many accounts of explanation (causal, deductive, analogical, model-based, unificationist, mechanistic, functional, probabilistic, and intentional-action), and good reasons to think that each of these forms of explanation is permissible in some contexts[7].

This is important, as concepts of causation that work at higher (e.g. macroscopic) levels of physical description do not tidily apply at the level of fundamental physics (Norton 2007), though certain models of non-causal explanation do (Thalos 2002). For instance, since Hempel (1965), many philosophers of science treat certain kinds of theoretical derivations as explanatory, and fundamental physics often trades in just these derivations (Cushing 1991). As such, while the philosophers discussed in this section consider the explanations they examine to be causal, it suffices for my purposes if the logical, probabilistic, and counterfactual dependence[8] relations invoked below are adequate for causal *or* non-causal explanation.

Kvanvig's Example certainly exhibits derivations that are at least candidates for explanation. Scientists typically derive probability distributions about an electron's position from its quantum state, i.e. the set of quantum numbers and the eigenfunction that characterize the possible states of the

[7] Many of the general reviews of the explanation literature discuss these different forms of explanation, e.g. (Cartwright 2004; Lipton 2004, Ch. 2; Lycan 2002; Salmon 1989; Thagard 1992, 118-130; Woodward 2002). Discussions about explanatory pluralism include (Achinstein 1983; Douglas 2009; Khalifa 2010; Risjord 2000; Ruben 1990; Thalos 2002; van Fraassen 1980).
[8] For a discussion of counterfactual, non-causal explanation see (Woodward 2003, 220-221).

quantum mechanical system of which the electron is a member. As such, there is not yet a reason to grant K2.

2.2.2. Indeterministic explanation

Given that Kvanvig's Example need not trade in causal explanation, Kvanvig's requirement about indeterminism becomes somewhat moot. While it is a near-tautology that uncaused events cannot be causally explained, uncaused events may very well admit to non-causal explanation. However, Kvanvig's remarks on this topic suggest misgivings about another notion of indeterminism that figures more prominently in the explanation literature. After presenting the core ideas of this latter notion of indeterministic explanation, I address Kvanvig's objection to it.

In the explanation literature, indeterminism simply means that some events are inherently chancy[9]. Causally, this does not mean that events are uncaused, but only that the same causes do not always produce the same effect. For example, a coin coming up heads is undoubtedly caused by a coin toss, but if coin tossing is an indeterminstic process, then an identical toss could have also caused the coin to land tails. Consequently, the core idea of indeterministic *explanation* is that "a factor *A* is explanatorily relevant to [an explanandum] *E* if *A* plays a non-eliminable role in determining the probability of *E*," where *A* exhausts the explanatorily relevant information (Hitchcock 1999, 587). Hereafter, I will use "indeterminism" in this sense.

[9] For a good review of the literature and the core ideas of indeterministic explanation, see (Glymour 2007). Glymour also argues that causal, indeterministic, contrastive explanations of equally probable outcomes are possible, but the details of his view are not necessary for what follows.

Many prominent theories of explanation of the past thirty years countenance indeterminism, e.g. (Humphreys 1989; Lewis 1986; Railton 1978, 1981; Salmon 1984; Woodward 2003)[10]. Peter Railton's Deductive Nomological Probabilistic (DNP) Model is one of the earliest and best-known expressions of indeterministic explanation[11]. Formally, Railton presents his model as:

(R1) $\forall t \forall x[F_{x,t} \rightarrow \text{Prob(G)}_{x,t} = p]$

"At any time, anything that is *F* has probability *p* to be *G.*"

Next, we adduce the relevant fact(s) about the case at hand, *e:*

(R2) $F_{e,to}$

"e is *F* at time t_o*,"*

and draw the obvious conclusion:

(R3) $\text{Prob(G)}_{e,to} = p$

"e has probability *p* to be G at time t_o*'"*

To which we add parenthetically, and according to how things turn out:

(R4) $(G_{e,to}/\neg G_{e,to})$

"(e did/did not become G at t_o*)."* (Railton 1978, 218: my numbering)

Here, the inference from (R1) and (R2) to (R3) must be sound, and the addendum (R4) must also be true. Railton also requires that (R1) refer to a

[10] Indeed, the only prominent outlier is (Kitcher 1989), who denies indeterminism on the grounds that explanations must be deductive arguments. As it turns out, determinism and an explanatory deductivism quite different than Kitcher's might be compatible (Glymour 2007). Moreover, Kitcher's deductivism has its own host of problems (Barnes 1992; Gijsbers 2007).

[11] I use Railton's account only because of its popularity and elegance, but should another account of explanatory indeterminism prove more satisfactory, my arguments about understanding would remain untouched.

"causal law," though for our purposes, any legitimate explanatory generalization will do.

Some of Kvanvig's remarks suggest he is skeptical of this kind of indeterminstic explanation. For instance, Kvanvig (2009a, 101) criticizes those who countenance explanations that "reify chance" into "a further explanatory or causal factor." Indeterminists take the connection between an event's probability to its actually happening to figure in an explanation, e.g. (R3) to (R4) in Railton's general schema. So, Kvanvig may be rejecting explanatory indeterminism outright.

Admittedly, the explanatory viability of such parenthetic addenda is not immediately intuitive, but there are two responses to Kvanvig's objection. First, indeterministic explanations are staples of scientific practice. For instance, Railton offers the following example:

(a) All nuclei of U^{238} have probability $(1 - \exp(-\lambda_{238} \bullet \theta))$ to emit an alpha-particle during any interval of length θ, unless subjected to environmental radiation.

(b) *u* was a nucleus of U^{238} at time *to,* and was subjected to no environmental radiation before or during the interval $t_o - (t_o + \theta)$.

(c) *u* had probability $(1 - \exp(-\lambda_{238} \bullet \theta))$ to emit an alpha-particle during the interval $t_o - (t_o + \theta)$…

(d) A parenthetic addendum to the effect that *u* did alpha-decay during the interval $t_o - (t_o + \theta)$ (Railton 1978, 214)

Other examples of indeterministic explanation pepper the natural and social sciences[12]. As a result, if Kvanvig is denying this kind of explanation *tout court*, he is at odds with scientific practice, home to some of our most exemplary feats of explanation and understanding.

Second, Kvanvig's objection does not acknowledge that indeterministic explanations also include theoretical statements like (R1) and (R2). Scientists do not baldly appeal to a mysterious explanatory factor called "chance," but link a single occurrence—such as an electron's trajectory—to well-defined physical propensities of a system, derivable from the theory (of which the quantum state is a part). Thus, the probabilities or "chances" that Kvanvig decries as non-explanatory are derived from theories that are undoubtedly explanatory.

Considerations of the counterfactuals operating in examples of indeterministic explanation highlight the explanatory import of these theoretical elements. For instance, in Railton's example, if *u* were not a ^{238}U nucleus (but rather, e.g. an ^{241}Am nucleus), the probability of alpha-decay would be different. Since these counterfactuals seem to track closely with our intuitions about explanatory relevance[13], there is no reason to resist this kind of explanation.

Importantly, a non-contrastive variant of Kvanvig's Example lends itself to indeterministic explanation. Derivations from quantum states can tell us why an electron had a probability *p* of being in a spatial region *x* at a given time interval *t*. Explanatory indeterminism then suggests that we can explain why the electron

[12] (Craver 2007) provides many neuroscientific examples; (Woodward 2003), examples from a wide variety of sciences.
[13] Importantly, these theoretical claims allow us to ascertain how the explanandum would have changed if the explanans had been different, an important feature of explanation (Woodward 2003).

was in *x* during *t*, simply by treating it as a "parenthetic addendum" to the derivation. As with Railton's example, the counterfactual holds: if the quantum state were different, then the probability of the electron being in a spatial region (e.g. "the left") would be different. Consequently, the indeterministic elements of this explanation do not lend credence to K2 in and of themselves.

2.2.3. Contrastive explanation

Thus, the orthodoxy on indeterministic explanation puts a good deal of pressure on Kvanvig to capitulate, minimally, that we can explain why an electron went left, even if they would agree with him that we can't explain why it went left rather than right. Indeed, despite the aforementioned authors' commitment to indeterministic non-contrastive explanation, all (save Woodward) think that "Contrastive Explanations Imply Determinism (CEID)" (Hitchcock 1999, 586). Kvanvig (2009a, 101-102) also endorses CEID:

> …the events in question are irreducibly indeterministic in such a way that there is no causal explanation as to why the actual events occurred rather than some other events.

If correct, CEID would advance Kvanvig's case significantly. I will first examine the justification for CEID, and then argue that the justification entails reductionism about objectual understanding. As such, either CEID is justified and reductionism is true, or CEID is unjustified, which also opens the door for reductionism by way of the Hidden Explanation Objection.

So why should we think that CEID is true? At the most basic level, contrastive explanation requires an explanans to be a 'difference-maker,' i.e. to

'discriminate' between the outcomes contrasted in an explanandum. For example, when we ask why Adam ate the apple rather than a candy bar, we are looking for something that favors Adam's eating the apple over the candy bar (e.g. Adam was on a diet).

However, we typically provide the same indeterministic (non-contrastive) explanation for different outcomes. Following Glymour (2007, 139), let us call this the thesis of *parity:* "one can [indeterministically] explain unlikely outcomes just as well as one can [indeterministically] explain their more probable alternatives." Essentially, parity arises because the same factors produce both a likely outcome and an unlikely one—that is the crux of indeterminism. For example, (R4) in Railton's DNP schema appears to clash with contrastivism, as it suggests that we should explain *e*'s *becoming G* or *e's not becoming G* in exactly the same way, i.e. by appeal to the derivation of (R3) from (R1) and (R2). As a result, there is nothing that can differentiate contrasted, indeterministic outcomes, giving credence to CEID.

Unfortunately for Kvanvig, parity strongly suggests reductionism. If parity holds, then the same information explains all of the possible outcomes of a given system. Thus, the same information (non-contrastively) explains why the electron went left, why it didn't go right, why it could have gone right, etc. However, since this is indeterministic, no additional information is relevant to the contrast. But if no additional information is relevant to the contrast, then there is no structural relationship that could underwrite even *objectual* understanding of it.

To illustrate this point, consider a fairly common presupposition in the indeterministic explanation literature:

> *A* will be said to be explanatorily relevant to *E* when P(*E*|*A & B*) ≠ P(*E*|B) (Hitchcock 1999, 587).

Here *B* refers to background conditions that are kept fixed in order to rule out spurious correlations. So, the only non-explanatory relationships left to grasp are either probabilistically irrelevant, or spurious correlations in which the background conditions *B* are not held fixed. Prima facie, even someone sympathetic to Kvanvig's position should view these probabilistic relations as implausible bases for objectual understanding.

In this case, we have an explanation of the electron's going left and also an explanation of its not going right (which, incidentally, share the same explanans), but nothing probabilistically relevant to the electron's going left rather than right can be grasped. Hence the contrast is understood neither objectually nor explanatorily, and everything in the vicinity of the contrast is understood explanatorily. Thus, while K2 would be true, Kvanvig's Example would only partake in explanatory understanding. So parity favors reductionism.

If, on the other hand, parity doesn't hold, then CEID lacks any obvious justification[14]. This paves the way for contrastive indeterministic explanations, such as Hitchcock's:

> *A* is explanatorily relevant to *E rather than F* when P*(E|(A & B) & (E∨F))* ≠ P*(E| B & (E∨F))* (Hitchcock 1999, 587)[15].

[14] For more on parity and/or contrastive indeterministic explanation, see (Glymour 2007; Hitchcock 1999; Strevens 2000).

Returning to Kvanvig's Example, since different quantum states (*A*) can change the probability of the electron's going left (*E*), even when it is presupposed that the electron went either left or right (*E∨F*), the electron's going left rather than right is explainable. In this case, K2 is false, as the only understanding in Kvanvig's Example is explanatory. Thus, regardless of one's stance on parity, CEID, or K2, reductionism appears to carry the day.

2.2.4. Equal probabilities

But suppose that Kvanvig denied CEID. His Example still might involve non-explanatory understanding, if, as he assumes, indeterministic contrastive explanations are possible only when the probabilities of the contrasted outcomes are different. Since Kvanvig's Example also assumes that the probability of the electron's going left is equal to that of its going right, indeterministic contrastive explanations are not possible in this particular case, even if they are possible elsewhere.

Such a position confuses the source of explanatory relevance or difference-making. An explanation needn't make the probabilities between contrasted outcomes different from each other; rather, these probabilities must be different than they would be had the explanans been different[16]. For example, suppose that my being aware of a sea bass dinner special yields a fifty-fifty chance that I choose that special—otherwise I order my old standby, eggplant. However, if I'm unaware of the special, then there is a negligible chance that I order sea bass.

[15] The disjunction (*E∨F)* is exclusive.
[16] While this is consistent with Hitchcock's view, (Woodward 2003) provides additional details.

Intuitively, my being aware of the special is explanatorily relevant to why I ordered sea bass rather than the eggplant[17].

As this example illustrates, Hitchcock's account of explanation, $P(E|(A \& B) \& (E \vee F)) \neq P(E|B \& (E \vee F))$, is consistent with equally probable outcomes, i.e. $P(E|(A \& B) \& (E \vee F)) = P(F|(A \& B) \& (E \vee F))$. Indeed, there is something odd about denying this. Suppose that we moved a polarizer in a continuous fashion and observed the corresponding changes in photon transmission and absorption. On Hitchcock's view, we can explain photon transmission and absorption for every conceivable orientation of the polarizer. By contrast, if there is something special about equal probabilities, the physical connection between the polarizer's orientation and the photon's behavior must be momentarily and miraculously interrupted precisely when the probability of transmission and absorption are identical, but then miraculously resumes as soon as those probabilities change. This is highly counterintuitive, to say the least. Thus, equally probable outcomes readily admit to contrastive explanation.

2.3. Discussion

Thus, nothing precludes the possibility of indeterministically explaining a contrast between two equally probable outcomes. This casts serious doubt on K2. However, there is a more general point here. These were very demanding constraints to put on an explanation, and yet, explanation emerged unfazed. This suggests that explanations pervade our understanding, as deterministic

[17] This is a variation on (Hitchcock 1999, 602-606)

explanations, non-contrastive explanations, and contrastive explanations of events with different probabilities are far less foreign than what we've been considering here. Additionally, philosophers of science would most likely find the denial of parity to be the most contentious move in my discussion, but as we saw, parity favors reductionism about objectual understanding. Consequently, even if the Hidden Explanation Objection falls short in Kvanvig's Example, non-explanatory understanding may still be very hard to come by.

Of course, this doesn't prove that all understanding is explanatory, but it shifts the burden of proof in that direction. For instance, formal disciplines such as mathematics are perhaps the most promising places to find logical and probabilistic relationships that are non-explanatory yet understanding-conferring. However, even here, explanation is not absent (Mancosu 2008). Hence, if the Hidden Explanation Objection is our only test, we should navigate the explanation literature with great care before reaching any definitive judgments for or against reductionism.

3. A proposed reduction

However, I believe that the Hidden Explanation Strategy is a special case of a much broader framework for a new reductionist program. This framework involves two key ideas. The first is a general methodological point about reduction (§3.1), while the second is a substantive assumption about the roles of logical and probabilistic relationships in explanatory understanding (§3.2). After presenting these ideas, I discuss how to test this reductionism (§3.3). While I provide reasons for thinking that reductionism passes these tests, I regard what

follows primarily as an invitation for subsequent discussions with antireductionists.

3.1. Fair comparisons

Early on, we noted that objectual understanding's grammatical form differed from that of explanatory understanding. In particular, objectual understanding could assume any grammatical object, while explanatory understanding only concerned embedded why-questions. Purely syntactical differences shouldn't impede the reductions I am suggesting; so two important points are in order.

First, understanding *why* is too narrow for capturing all explanatory understanding. Answers to some how-questions, e.g. "How do amoebae reproduce?" are also explanations. Thus, explanatory understanding can take either embedded why- or how-questions as its object[18].

Second, and more importantly, the embedded question that characterizes a case of explanatory understanding should be matched with its proper correlate in objectual understanding. Call this the *Fair Comparison Requirement*. In our earlier example, we have:

(A) Leticia (objectually) understands the occurrence of the Argentinean economic crisis of 2001-2002.

(B) Adriana (explanatorily) understands how/why the Argentinean economic crisis of 2001-2002 occurred.

That is as fair a comparison as one can find. However, antireductionists are frequently tempted to compare (B) with something like:

[18] Note that there is another kind of understanding-how that is of a practical variety, e.g. Jimi understands how to play guitar. This is clearly not explanatory.

(A') Leticia (objectually) understands Argentinean economics.

It is no surprise that (A') does not reduce to (B), as it concerns a different topic. If (A') were the kind of objectual understanding that interested us, the proper comparison would be to something resembling:

(B') Adriana (explanatorily) understands how/why Argentinean economics does what it does/is what it is.

While a more thoroughgoing syntactical analysis would be desirable, I trust that what I am requiring is clear enough: comparing objectual apples with explanatory oranges makes for poor defenses of antireductionism. By analogy, consider that a particular column of air A_1 reduces to a particular set of molecules M_1. It is no criticism of reductionism to say that another column of air A_2 does not reduce to the same set of molecules M_1. This is true even when A_2 contains A_1, as antireductionists might say about (A') and (B). What goes for air columns also goes for understanding.

3.2. Explanatory roles

The core of my reductionism involves a strategy for subverting logical and probabilistic relationships to explanation. The Hidden Explanation Objection is a special case of that strategy, but there is a more general approach lurking beneath it. Let *Rab* denote that two items of information *a* and *b* stand in structural relationship *R*. Then *Rab* plays an explanatory role in *S's* understanding *how/why p*[19] if *S's* grasping that *Rab* enables *S* to either:

[19] For economy of prose, I will sometimes say that the relationship or that a relatum plays an explanatory role, rather than that the two informational items' standing in the relationship plays that role.

(1) Correctly explain *p*; or

(2) Identify the marks that make some item of information *q* a better explanation of *p* than another rival explanation of *p*. [20]

Now, let us assume that grasping logical and probabilistic relationships contributes to understanding only to the extent that these relationships play an explanatory role. Call this the *Explanatory Role Assumption* (ERA).

Let me offer four examples of how logical and probabilistic relationships can play explanatory roles, and briefly suggest how they fit within my reductionist program. First, logical and probabilistic relationships are frequently explanatory, as we saw in the Hidden Explanation Objection. Clearly, in these cases, they are just explanatory relationships, and thus amenable to reduction.

Second, logical and probabilistic information may be either an explanans or an explanandum. Railton's DNP model illustrates how such information could figure in an explanans. However, pieces of logical and probabilistic information also frequently serve as explananda. On the probabilistic side, information about correlations is frequently explained. For example, the fact that smoking causes cancer explains why smoking is correlated with cancer. Similarly, Kepler's laws contain logical relationships (primarily the identity relation), and were explained by Newtonian mechanics. Since explanantia and explananda are essential elements of an explanation, their including logical and probabilistic relationships does not license antireductionism.

[20] See (Khalifa forthcoming) for more on these two abilities' relevance to an analysis of understanding.

Third, even when logical and probabilistic relationships are not "directly" explanatory, they may still be justifying parts of an explanation (i.e. the explanans, explanandum, or the fact that the two stand in a given explanatory relation). For example, suppose that Leticia grasps logical and probabilistic relationships that give her better justification for the same explanation that is at Adriana's disposal. Then Leticia can grasp things about the goodness of that explanation that are opaque to Adriana. In this case, Leticia simply has better explanatory understanding than Adriana, but she need not have a different, irreducible kind of objectual understanding.

Finally, logical and probabilistic relations can facilitate correct explanations by specifying the *presuppositions* of a correct explanation (Garfinkel 1981; Hitchcock 1999; Khalifa 2010; Risjord 2000; Sober 1986). For instance, my arm bumping the inkwell explains why it spilled, and the inkwell's spilling presupposes (e.g., through logical entailment and auxiliary assumptions) that an object is extended in space. The relationship between the inkwell's spilling and the presupposition is not explanatory, yet without the presupposition, correctly explaining it would be difficult if not impossible (e.g. try explaining why the inkwell spilled if it could have been a one-dimensional object). This also is compatible with reductionism, as grasping an explanation's presuppositions is simply grasping further aspects of the explanation.

Importantly, ERA does not entail that everything within one's grasp must be explained. If one likes, certain information helps to *provide* (explanatory) understanding of something else, even if it is not itself understood. In the previous example, one may not have the foggiest idea why objects are extended in

space (depending on how one reads the explanatory demand here, it may require appeal to Big Bang cosmology). In this case, this presupposition would be unexplained but is indispensable to grasping an explanation of the inkwell's spilling. Thus, unexplained information that plays an explanatory role is no threat to the reductionism presented here. This proves crucial in what follows.

3.3. Successful criteria

My reductionist picture thus amounts to the following: wherever there is a fair comparison between objectual and explanatory understanding, logical and probabilistic relationships provide understanding only inasmuch as they are playing explanatory roles. If true, objectual understanding is reducible to explanatory understanding. Let me briefly present how this works by discussing potential responses to two kinds of antireductionist arguments.

3.3.1. Demanding antireductionism

One kind of antireductionist might argue that explanatory understanding does not always *suffice* to provide objectual understanding. Even assuming we have a fair comparison, these *demanding antireductionists* claim that objectual understanding involves *more* than explanatory understanding. Since these two forms of understanding only differ with respect to the structural relationships involved, these antireductionists take objectual understanding to involve grasping more of these relationships than would be involved in explanatory understanding. For example, the demanding antireductionist might cite the fact that Leticia's understanding is irreducibly richer than Adriana's if, in addition to

their shared explanatory information, Leticia can grasp additional logical and probabilistic relationships that are opaque to Adriana.

This poses no threat to reductionism for two reasons. First, demanding antireductionism assumes that explanatory understanding involves grasping *only* explanatory relationships. However, this does not follow from our discussion in §1.1. There, we claimed that, of the relationships grasped, at least one had to be explanatory. However, this is compatible with other structural relationships being non-explanatory.

If this is correct, we can follow Kvanvig in treating Leticia's understanding as *more coherent* than Adriana's[21]. Since demanding antireductionism also assumes that Leticia and Adriana grasp a common explanatory relation, this suggests that Leticia's *explanation* is more coherent than Adriana's. But then demanding antireductionism could just as well be called "demanding reductionism," as the only difference is the quality of Leticia's explanation.

Perhaps there are counterexamples to this last claim, but this brings us to our second point: demanding antireductionists are constrained by fair comparison. I submit that if an explanation figuring in the demanding antireductionist's example of objectual understanding is no more coherent than it would be in a case of high-grade explanatory understanding, we have reason to question whether there is a fair comparison.

For example, suppose that Leticia has an impressive understanding of Argentinean market behavior in 1999, but she does not grasp how those markets

[21] E.g. "[U]nderstanding requires… an internal grasping or appreciation of how the various elements in a body of information are related to each other in terms of explanatory, logical, probabilistic, and other kinds of relations that coherentists have thought constitutive of justification (Kvanvig 2003, 192-193).

affected the crisis of 2001-2002. Furthermore, her explanation of the latter is identical to Adriana's. There may be some sense in which her understanding of the 1999 markets is relevant to the 2001-2002 economic crisis, but we could just as easily gloss this as two unrelated feats of explanatory understanding, i.e. Leticia understands how/why the economic crisis of 2001-2002 occurred and she understands how/why the Argentinean markets of 1999 behaved as they did.

If this is correct, then fair comparisons provide compelling support for ERA. If the demanding antireductionist's added structural relationships don't play a role in the explanation, then they concern a different topic of understanding, and are thereby irrelevant to the reduction at hand. If, on the other hand, these relationships do play an explanatory role, demanding antireductionism amounts to the claim that objectual understanding is just having more coherent (and presumably better) explanatory understanding. But that is reductionism. So my view appears immune to demanding antireductionist arguments.

3.3.2. Humbling antireductionism

However, not all antireductionists are demanding. Some preach *humility*, arguing that explanation is not *necessary* for understanding. On this view, a person may grasp only non-explanatory relationships and still have objectual understanding while lacking explanatory understanding. For argument's sake, suppose that Kvanvig's Example did not involve a hidden explanation. The humbling antireductionist could then offer the following fair comparison:

(C) Paul understands the electron's going left rather than right.

(D) Paul understands how/why the electron went left rather than right.

Since we are assuming that that K2 is true, (D) would be false. However, because Paul grasps non-explanatory relationships linking information about the electron's going left rather than right to other information, the humbling antireductionist claims that she possesses non-explanatory, objectual understanding.

By contrast, my reductionism claims that if (D) is false, then so is (C). So the central point is whether the non-explanatory relationships that Paul grasps provide understanding of the electron's trajectory. ERA indicates how reductionists can still say that these relationships provide explanatory understanding of *some other phenomenon*, even if they do not provide this understanding with respect to the electron's going left rather than right. This would allow us to maintain a strong reductionist position, as there would be no need to interpret (C) as true, much less involving an irreducible notion of objectual understanding.

To see this, suppose we assume, as we did in the Failed Criterion Objection, that what Paul really understands is a neighboring explanation of a system:

(E) Paul understands how/why system *S* does what it does/is what it is.

Now, one thing this system "does" is to render deterministic explanations of electron trajectories impossible. So a person who understands how the system works grasps the following "meta-explanation" (e.g., by way of the Uncertainty Principle):

(F) Paul understands why it is impossible to deterministically explain how/why electrons in *S* go in one direction rather than another.

Now assuming that Paul has robust understanding of why the system does what it does, he grasps many presuppositions of (F), e.g.

(G) Paul grasps that *if electron e in S went left rather than right, it cannot be deterministically explained*.

In (G), Paul grasps a non-explanatory, logical relationship. While the humbling antireductionist cites this as evidence for (C), nothing in the story that brought us from (E) to (G) is inconsistent with reductionism. Per ERA, the understanding in (G) is exhausted by its role as a presupposition of the explanation in (F), i.e. the explanation of why deterministically explaining electron trajectories is impossible. (F), in turn, is simply part of understanding how/why the system does what it does (E). As such, there is no sense in which *e's* going left rather than right is explained, yet there is also no sense in which it is involved in non-explanatory understanding.

At best, this shows that reductionists have plausible ways of dealing with humbling antireductionists. However, I currently lack a decisive argument against the humbling antireductionist that would not simply express a difference of intuitions. For instance, I'm inclined to think that a person who only grasps that *A is correlated with B* lacks understanding of both *A* and *B*, while a person who can explain *A* with *B* but is unaware of this correlation nevertheless understands how *A* could have happened. Similarly, I think that proficient deductive and probabilistic reasoners who lack explanations might be able to *classify* and *describe*—but do not *understand*—a body of information. Perhaps

these intuitions are more powerful and pervasive than I think, which would lend further credence to reductionism. However, I would be unsurprised if humbling antireductionists had different intuitions about these matters.

That being said, the humbling antireductionist can be pressed further. The example involving Paul is simply a generalized version of the Hidden Explanation Objection—we might call it the Hidden Explanatory Role Objection (HERO). Whereas the Hidden Explanation Objection entailed that structural relationships offered as counterexamples to reductionism had to be *explanatory*, HERO only requires these relationships play an *explanatory role*, such as the presupposition (G) in the example just rehearsed.

When we pair HERO with a healthy explanatory pluralism, the space for justifying humbling antireductionism shrinks considerably. The only logical and probabilistic relationships available are those playing none of the explanatory roles discussed in §3.2 (explanatory relation, explanandum, explanans, explanatory justification, presupposition), wherein explanations can assume any of the forms discussed in §2.2 (causal, deductive, analogical, model-based, unificationist, mechanistic, functional, probabilistic, intentional-action, mathematical). Moreover, there is no guarantee that either of these lists is complete.

Nor is it obvious that the logical and probabilistic relationships that survive HERO provide understanding. For instance, in §2.2.3, we saw that the details of indeterministic explanation, specifically the parity thesis, precluded the possibility of additional probabilistically relevant information. Consequently, the

humbling antireductionist could only appeal to probabilistically irrelevant information, which is a fallow source of understanding, to say the least.

In summary, we have explored two ways to answer the Failed Criterion Objection, and neither casts a favorable light on antireductionism. The first was demanding antireductionism, which only appears to succeed if our requirements for a fair comparison are violated. The second was humbling antireductionism, which shoulders a substantial burden of proof that has not been met.

4. Conclusion

To summarize, Kvanvig distinguished between objectual and explanatory understanding. The central difference between them concerned the importance of explanation in providing understanding. Kvanvig's Example, a putative case of objectual understanding without explanation, failed to establish antireductionism for two reasons. First, per the Failed Criterion Objection, Kvanvig's test for reducibility was inadequate. Second, per the Hidden Explanation Objection, Kvanvig's alleged example of an unexplained fact might well be explainable, or at the very least, amenable to reduction.

A more developed framework for reduction was presented. Its core ideas were the Fair Comparison Requirement, which specifies the relata of the reduction, and the Explanatory Role Assumption, which subverts logical and probabilistic relationships to explanatory ones in the context of understanding. We then showed how these ideas provided principled answers to demanding and humbling antireductionism.

There are many loose ends. There may be other forms of reduction, trickier examples, and some lacunae in the reductionist picture sketched here. But in pursuing all of these topics, nothing appears more fruitful than a considered judgment of where the explanation literature comes up short in illuminating our concept of understanding—if it comes up short at all.

Achinstein, P. (1983), *The nature of explanation*. New York: Oxford University Press.
Barnes, E. C. (1992), "Explanatory Unification and the Problem of Asymmetry", *Philosophy of Science* 59 (4):558-571.
Cartwright, N. (2004), "From causation to explanation and back", in B. Leiter (ed.), *The future for philosophy*, Oxford: Oxford University Press, 230-245.
Craver, C. (2007), *Explaining the brain: mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.
Cushing, J. T. (1991), "Quantum theory and explanatory discourse: endgame for understanding?", *Philosophy of Science* 58 (3):337-358.
de Regt, H.W. (2004), "Discussion note: making sense of understanding", *Philosophy of Science* 71:98-109.
——— (2009a), "The epistemic value of understanding", *Philosophy of Science* 76 (5):585-597.
de Regt, H.W., and D. Dieks (2005), "A contextual approach to scientific understanding", *Synthese* 144 (1):137-170.
de Regt, H.W., S. Leonelli, and K. Eigner, eds. (2009), *Scientific understanding: philosophical perspectives*. Pittsburgh, Pa.: University of Pittsburgh Press.
Douglas, H. E. (2009), "Reintroducing Prediction to Explanation", *Philosophy of Science* 76 (4):444-463.
Elgin, C. (2006), "From knowledge to understanding", in S. Hetherington (ed.), *Epistemology futures*, Oxford: Oxford University Press, 199-215.
——— (2007), "Understanding and the facts", *Philosophical Studies* 132 (1):33-42.
——— (2009), "Is understanding factive?", in A. Haddock, A. Millar and D. Pritchard (eds.), *Epistemic Value*, Oxford: Oxford University Press, 322-330.
Garfinkel, A. (1981), *Forms of explanation: rethinking the questions in social theory*. New Haven: Yale University Press.
Gijsbers, Victor (2007), "Why Unification Is Neither Necessary Nor Sufficient for Explanation", *Philosophy of Science* 74 (4):481-500.
Glymour, B. (2007), "In defence of explanatory deductivism", in J.K. Campbell, M. O'Rourke and H. Silverstein (eds.), *Causation and Explanation*, Cambridge, MA: MIT Press, 133-154.
Grimm, S. R. (2006), "Is understanding a species of knowledge?", *British journal for the philosophy of science* 57 (3):515-535.

——— (2008), "Explanatory inquiry and the need for explanation", *British journal for the philosophy of science* 59 (3):481-497.
——— (forthcoming), "The goal of understanding", *Studies in the history and philosophy of science*.
Gross, S. (2005), "Linguistic understanding and belief", *Mind* 114 (453):61-66.
Hempel, C. G. (1965), *Aspects of scientific explanation, and other essays in the philosophy of science*. New York: Free Press.
Hitchcock, C.R. (1999), "Contrastive explanation and the demons of determinism", *British journal for the philosophy of science* 50 (4):585-612.
Humphreys, P. (1989), *The chances of explanation*. Princeton: Princeton University Press.
Khalifa, K. (2010), "Contrastive explanations as social accounts", *Social Epistemology* 24 (4):265-286.
——— (forthcoming), "Understanding, knowledge, and scientific antirealism", *Grazer Philosophische Studien* 83.
Kitcher, P. (1989), "Explanatory unification and the causal structure of the world", in Philip Kitcher and Wesley C. Salmon (eds.), *Scientific Explanation*, Minneapolis: University of Minnesota Press, 410-506.
Kvanvig, J. L. (2003), *The value of knowledge and the pursuit of understanding*. Cambridge: Cambridge University Press.
——— (2009a), "The value of understanding", in A. Haddock, A. Millar and D. Pritchard (eds.), *Epistemic value*, Oxford: Oxford Unviersity Press, 95-111.
——— (2009b), "Responses to critics", in A. Haddock, A. Millar and D. Pritchard (eds.), *Epistemic value*, Oxford: Oxford University Press, 339-352.
Lacey, H. (1999), *Is science value free?: values and scientific understanding*. London: Routledge.
Lewis, D. (1986), "Causal Explanation", in, *Philosophical Papers*, Oxford: Oxford University Press, 214–240.
Lipton, P. (2004), *Inference to the best explanation*. 2nd ed. New York: Routledge. Original edition, 1991.
Longworth, G. (2008), "Linguistic understanding and knowledge", *Nous* 42:50-79.
Lycan, W. G. (2002), "Explanation and epistemology", in Paul K. Moser (ed.), *The Oxford Handbook of Epistemology*, Oxford: Oxford University Press, 408-433.
Mancosu, P. (2008), "Mathematical Explanation: Why it Matters", in P. Mancosu (ed.), *The philosophy of mathematical practice*, Oxford: Oxford University Press, 134-151.
Norton, J. D. (2007), "Causation as folk science", in H. Price and R. Corry (eds.), *Causation, physics, and the constitution of reality: Russell's republic revisited*, Oxford: Oxford University Press, 11-44.
Pettit, D. (2002), "Why knowledge is unnecessary for understanding language", *Mind* 111 (443):519.
Pritchard, D. (2008), "Knowing the answer, understanding, and epistemic value", *Grazer Philosophische Studien* 77:325-339.

——— (2009), "Knowledge, understanding, and epistemic value", in A. O'Hear (ed.), *Epistemology*, Cambridge: Cambridge University Press, 19-43.
——— (2010), "Knowledge and understanding", in D. Pritchard, A. Millar and A. Haddock (eds.), *The nature and value of knowledge: three investigations*, Oxford: Oxford University Press, 3-90.
Railton, P. (1978), "A deductive-nomological model of probabilistic explanation", *Philosophy of Science* 45 (2):206-226.
——— (1981), "Probability, explanation, and information", *Synthese* 48:233-256.
Riggs, W. (2003), "Understanding 'virtue' and the virtue of understanding", in M. DePaul and L. Zagzebski (eds.), *Intellectual virtue: perspectives from ethics and epistemology*, Oxford: Oxford University Press, 203-226.
——— (2009), "Understanding, knowledge, and the meno requirement", in A. Haddock, A. Millar and D. Pritchard (eds.), *Epistemic value*, Oxford: Oxford University Press, 331-338.
Risjord, M. (2000), *Woodcutters and witchcraft: rationality and interpretive change in the social sciences*. Albany: State University of New York Press.
Ruben, D.H. (1990), *Explaining explanation*. New York: Routledge.
Salmon, W. (1984), *Scientific explanation and the causal structure of the world*. Princeton, N.J.: Princeton University Press.
——— (1989), "Four decades of scientific explanation", in Philip Kitcher and Wesley Salmon (eds.), *Scientific Explanation*, Minneapolis: University of Minnesota Press, 3-219.
Sober, E. (1986), "Explanatory presupposition", *Australasian journal of philosophy* 64:143-149.
Strevens, Michael (2000), "Do Large Probabilities Explain Better?", *Philosophy of Science* 67 (3):366-390.
Thagard, P. (1992), *Conceptual revolutions*. Princeton: Princeton University Press.
Thalos, M. (2002), "Explanation is a genus: an essay on the varieties of scientific explanation", *Synthese* 130 (3):317-354.
Trout, J.D. (2002), "Scientific explanation and the sense of understanding", *Philosophy of Science* 69:212-233.
——— (2005), "Paying the price for a theory of explanation: de Regt's discussion of Trout", *Philosophy of Science* 72:198-208.
——— (2007), "The psychology of scientific explanation", *Philosophy Compass* 2/3:564–591.
van Fraassen, B. C. (1980), *The scientific image*. New York: Clarendon Press.
Woodward, J. (2002), "Explanation", in Peter Machamer and Michael Silberstein (eds.), *Blackwell Guide to the Philosophy of Science* Malden: Blackwell, 37-54.
——— (2003), *Making things happen: a theory of causal explanation*. New York: Oxford University Press.
Zagzebski, L. (2001), "Recovering Understanding", in M. Steup (ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, Oxford: Oxford University Press, 235-252.