Depth and Deference: When and Why We Attribute Understanding

§1 Introduction

Epistemology has always been of central concern to philosophers: assumptions about how we can know and what we can know are arguably at the foundation of any further inquiry. While the particulars of contemporary epistemology would be foreign to most people outside of academic philosophy, epistemic concepts are still central to children's development and to everyday reasoning (e.g., Kuhn, Cheney, & Weinstock, 2000; Shatz, Wellman, & Spilber 1983). In the present paper, we aim to shed light on the relationship between contemporary epistemology and "folk" epistemology, and in particular to begin to characterize folk attributions of understanding as distinct from folk attributions of knowledge. Our central motivating question is: why do we attribute propositional understanding (e.g., 'Lisa understands why P happened') to some people but not to others? This question can be decomposed into two elements. First,

Basis Question: On what *basis* do we attribute understanding?

That is, what are the conditions under which it can be felicitously asserted that someone understands? Second,

Purpose Question: For what *purpose* do we attribute understanding?

What (potentially unique) epistemic aim are we trying to satisfy when we divide the world into understanders and non-understanders?

The answers to these two questions should go together—the basis for our understanding attributions should enable them to serve the role we want played when we pick out understanders. We propose a simple and intuitive two-pronged answer: identifying understanders lets us pick out people whose beliefs attain a certain degree of explanatory depth (=basis), *so that* we then know to consult those people on related problems (=purpose). This tells us something about how we attribute mental states to other people, and what practical purposes might be advanced by doing so. Moreover, we will argue that because deference plays a crucial role in many aspects of everyday reasoning (e.g., Hofer & Pintrich 1997, Keil 2003), the fact that we use understanding attributions to demarcate experts reveals a potential mechanism for achieving our epistemic aims in many domains.

To support our proposal, we report a series of experiments contrasting people's use of the locution 'X understands why P' with their use of the locution 'X knows why P'. In much common parlance, 'understands' and 'knows' are used interchangeably. Nonetheless, we think there is a great deal to be learned about the circumstances in which the usages do diverge and the differing epistemic expectations people have in such circumstances. We will argue that understanding-why attributions differ from knowledge-why attributions in that people have higher demands of explanatory depth before attributing understanding, and that this difference is related to speakers' interest in tracking who would be a valuable expert to consult about a particular field of inquiry. Since people with deeper explanations could be expected to successfully answer a more robust range of questions about a given phenomenon, the answers to the basis and purpose question thus seem to align. In this way, we

establish that understanding-why is a stronger epistemic state than mere knowledge-why, at least with regard to the dimension of explanatory depth. More generally, examining systematic patterns in our attributions of epistemic states (understands, knows, believes, etc.) should help reveal our diverse epistemic aims as cognitive agents in a complex world, and how we go about achieving them.

First, a disclaimer: We are looking entirely at the connection between understanding-why and knowledge-why, as opposed to understanding-of something ("objectual understanding"), such as a theory or a work of art (e.g., "understanding quantum mechanics" or "understanding the Mona Lisa"). To the extent that attributions of understanding-of are equivalent or reducible to attributions of understanding-why, many of our results might be expected to translate to the objectual case. If understanding-why is of a fundamentally different kind from understanding-of, then it seems reasonable to explore the *Purpose* and *Basis* questions with respect to understanding-why specifically. (For a more complete examination of objectual understanding, see Wilkenfeld 2013.)

In the next section, we look at philosophical arguments for why we should expect understanding to require greater depth than knowledge, but also explain why this link might not be as obvious as it seems. In §3, we explore some of the psychology of epistemology, with particular attention paid to the role of deference to expertise. In §4, we present data that suggest that people do differentially withhold understanding attributions on the basis of lack of depth, and that this difference recedes as depth is increased. In §5, we demonstrate this phenomenon at a finer grain and connect it to judgments about deference within and across domains.

§2 Understanding is Deeper than Knowledge

In this section, we argue on philosophical grounds that understanding is deeper than knowledge. Following Rosenberg (1981), we contend that understanding attributions are used for the purpose of identifying and demarcating experts to whom one should defer regarding both practical and theoretical questions. As a result, understanders are required to have a certain degree of depth to their beliefs, beyond merely assenting to the relevant explanatory content.

At first, it might seem trivial that there are instances of knowledge that are not instances of understanding. To borrow an example from Sosa (2000), there are innumerable truths about the precise relation connecting any two grains of sand in a desert. These facts could all be known, but it seems doubtful that such knowledge could ever amount to understanding.

This conclusion would be too hasty, however. As Brogaard ("I think. Therefore I understand."Unpublished manuscript, 2005) has argued, many apparent cases of understanding outstripping knowledge do not keep the object of understanding and knowledge constant. For example, in the sand case, we can know-that the sand is in a particular configuration, but the apparently obvious intuition that this is distinct from understanding illicitly presupposes that the sort of understanding in question is understanding-why. If we stick to the [*epistemic-state]-that* locution, it is less obvious that one could not have understanding-that the sand is in some particular configuration. Granted, we do not seem to have understanding-*why*, but we do not have knowledge-*why* either. Brogaard makes a similar point regarding the trap of comparing understanding-of with knowledge-why—obviously, to really

establish an asymmetry between knowledge and understanding, one must show that the epistemic states have different properties when their objects are held constant. (See Wilkenfeld, Plunkett, & Lombrozo, in prep, for relevant discussion, including empirical exploration of whether people really fall prey to the confusion Brogaard suggests.)

Consistent with our proposal, there has also been a suggestion in the literature that one should be able to differentiate understanding-why from knowledge-why in that the latter is attainable in isolation, whereas the former requires some degree of explanatory depth. Pritchard (2009, p. 38) asserts:

> "…it is possible to know why one's house has burned down (and indeed know that it burned down because of faulty wiring), even though one does not understand why one's house burned down…Suppose that I understand why my house burned down, know why it burned down, and also know that it burned down because of faulty wiring. Imagine further that my young son asks me why his house burned down and I tell him. He has no conception of how faulty wiring might cause a fire, so we could hardly imagine that merely knowing this much suffices to afford him understanding of why his house burned down. Nevertheless, he surely does know that his house burned down because of faulty wiring, and thus also knows why his house burned down. Indeed, we can imagine a teacher asking my son if he knows why his house burned down and him telling the teacher the reason. If asked by a second teacher if my son knew why his house burned down, we could then imagine the first teacher saying that he did. So, it seems, one can… [have] knowledge without the corresponding understanding."

This passage highlights a level of explanatory depth at which it would be appropriate to attribute outright knowledge-why but not outright understanding-why (as opposed to graded or explicitly relativized attributions). However, as Christoph Kelp (2015) persuasively argues, the level at which it becomes felicitous to say that someone does or does not understand simpliciter will involve a threshold that shifts on the basis of context. Pritchard's deeper insight, beyond a mere observation about when someone does or does not understand-why, is that at least one relevant *dimension* according to which we evaluate understanding-why as opposed to knowledge-why is that of explanatory depth. This suggests a hypothesis, which we call *Depth Difference*: One's level of understanding-why tracks explanatory depth more closely than does one's level of knowledge-why.

Depth Difference also complements another point by Kelp (2015), who argues that ideal understanding is maximally comprehensive and well-connected knowledge. On this view, maximal understanding would necessarily be deep. Kelp contends that outright understanding attributions are proper when the degree to which an agent approximates ideal understanding suffices for contextually-relevant tasks. Our contention that the epistemic aim of understanding is to pick out people to whom we should defer does not precisely align with the claim that (outright) understanders are those who could perform a task we care about, but it is a close analogue to it. In particular, if *Depth Difference* predicts people's patterns of understanding attributions, that provides evidence for, and more detail regarding, Kelp's contention that we deem people 'understanders' when they have the sort of information we care about in a particular context.

A point in favor of *Depth Difference*, aside from the intuitive pull of cases such as Pritchard's, is the close conceptual link between understanding and explanation. As discussed at various points in the

literature (Friedman, 1974; Kim 1994), even theorists of explanation as anti-psychologistic as Hempel (1965) would have granted that explanations tend to produce understanding. On the extreme end, some (Scriven, 1962; Wilkenfeld, 2014) have even argued that it is constitutive of being an explanation that something play the right role in generating understanding. Thus, to the extent that depth is an important attribute of explanations, we might expect it to play a disproportionate role in evaluating and individuating understanding-why (even relative to knowledge-why). In fact, depth does play a large evaluative role in theories of explanation, as it is presupposed by the mechanistic accounts that demand complete productive intelligibility (Machamer, Darden, & Craver, 2001) and those that rate accuracy in terms of grain of description (Weiskopf, 2011). More generally, the importance of depth (specifically as it relates to causal power and abstraction) is a centerpiece of Strevens's (2008) exploration of explanation (and indeed even of causation itself).

In sum, *Depth Difference* predicts that people will be more willing to attribute knowledge-why than understanding-why in cases where someone has a true explanatory belief that is nevertheless fairly shallow. It also hints at a possible epistemic aim to our practice of attributing understanding: understanding attributions play a functional role in identifying experts who should be heeded with respect to the general field in question. This would lead to a second prediction, that people's attributions of understanding closely track their judgments concerning whom they would be likely to go to as experts. In the four experiments we go on to report, these predictions are borne out.

§3 On the Psychology of Epistemology

Within the psychology literature, there has been some exploration of people's concepts of knowledge and how those concepts change over time. Shatz, Wellman, and Spilber (1983) studied the early-childhood development of epistemic concepts, and argue that those concepts form an important component of children's conceptual scheme just before age 3. Suggestively, they also find some evidence (p. 318) that children form beliefs about others' mental states at about the same time they form metacognitive beliefs about their own (a finding demonstrated for a larger group of children in Bartsch & Wellman, 1995). Thus, from our first uses of mental state concepts such as *knows,* we apply them to other people. This suggests that the sort of attribution task we ask our participants to engage in (rating others' understanding and knowledge) is part and parcel with our deployment of mental concepts generally and theory of mind more broadly.

Other work has investigated how our relationship to experts develops throughout young adulthood. Hofer and Pintrich (1997) discuss numerous developmental models that detail how people's (and particularly students') approaches to knowledge vary as a function of time and education. Importantly for our purposes, all of these models suggest that people originally approach knowledge by way of deference to authority figures (in alignment with findings of Koenig and Harris, 2005, who explore children's ability to identify and defer to experts), then go through a stage in which they reject deference altogether, then (sometimes) return to a stance of selective deference. People's relation to experts is considered by many researchers to be one of the, if not the, most telltale markers of individual epistemological style (Hofer & Pintrich 1997). Kuhn, Cheney, and Weinstock (2000) add that we are slower to put down, and quicker to pick back up, objectivist approaches (which tacitly acknowledge some people's position as experts) in some domains over others (e.g., scientific claims versus claims of beauty).

Another viewpoint on the importance of deference to pursuing our epistemic aims is expressed by Frank Keil (2003) in his discussion of the "division of cognitive labor." Keil looks at the ways we form

detailed views: by filling in skeletal theories that spell out very broad generalities about what sorts of causal relationships govern specific domains. One important mechanism for this filling-out process involves our reliance on other members of our community (in a manner analogous to Putnam's contention that we divide linguistic labor by letting experts fix the extension of technical terms). No person, or even scientist, can be expected to know all the details of how something works, and so we *must* defer to others if we are to make any headway in learning about the world. Keil, Stein, Webb, Billings, and Rozenblit (2008) also suggest that deference is central to knowledge; they go one step further in spelling out the kinds of people we pick out as useful experts on a relevant topic, demonstrating that people group potential informants along the lines of traditional academic disciplines and are more likely to do so the more abstract the problem is.

All of this research suggests a question that our work aims to partially answer—who are the experts, and how do we mark them out as deserving of our attention? This research additionally underscores the important role of deference in our epistemological lives. Most theories explicitly tie such deference to *knowledge*. However, none of the studies looked for a distinction between *knowledge* and *understanding*, and so it would be surprising if the theories were sensitive to such a distinction. There is also the occasional hint that understanding and knowledge are being lumped together, as when Kuhn and Park (2005) describe the epistemological progression not in terms of knowledge, but in terms of "enhancing individual and collective understanding" (119). Our results, which more tightly link deference with understanding, are not in opposition to earlier findings, but indicate a possible refinement.

§4 The Data on Depth

Although Pritchard's case has some intuitive pull for philosophers, it is not uncommon within epistemology for seemingly-stable philosophical intuitions to diverge from so-called "folk" intuitions (e.g., Weinberg et al 2001, Starmans & Friedmans 2012, etc.). This risk is doubly present in the case of understanding, as Brogaard provides reason to doubt the universality of Pritchard's intuition even among philosophers. In this section we review the beginnings of the empirical evidence in favor of *Depth Difference*.

Experiment 1

We began by testing for potential differences in attributions of knowledge and understanding as a function of explanatory depth. Based on prior work (Wilkenfeld, Plunkett, & Lombrozo, in preparation.), we had good reason to expect that complete understanding and complete knowledge hang together very tightly in folk attributions. It's at intermediate levels of depth that we have an opportunity to observe attribtuions of understanding and knowledge to diverge. Accordingly, we contrasted a case of extremely *shallow* belief with a case of *no belief* about an explanatory connection. Cases of the former type are close analogs of Pritchard's original faulty wiring case above.

Method

Participants One-hundred-forty-two adults (69 female, mean age 36) were recruited through the Amazon Mechanical Turk marketplace (MTurk) and participated in exchange for monetary compensation. In all experiments, participation was restricted to users with an IP address within the United States and an approval rating of at least 95% based on at least 50 previous tasks. An additional

188 participants were excluded prior to analysis for failing to consent, failing to complete the experiment, or giving an incorrect response to one of the reading comprehension questions.

Materials and Procedure At the beginning of the experiment, participants were randomly assigned to read one of two vignettes describing a person, "Richard," and the extent and accuracy of his beliefs about certain features of the observed orbit of Mercury.

In the *shallow* condition, participants read:

> Richard has been taught a few things about astronomy and astrophysics. For example, if asked, he will tell you a few true things about the orbit of Mercury. He can accurately report Mercury's mass, volume, and average distance from the sun, and that Mercury's apparent orbit has some oddities.

> Richard can also tell you a few things about the theory of general relativity. One thing he learned is that general relativity explains some oddities in the observed orbit of Mercury, but he would not be able to answer any follow-up questions about that. Everything he says about Mercury and general relativity is true, though. Other than the fact that he's been told that general relativity explains the observed oddities of the orbit, he does not see any connections between the things he says about Mercury and about general relativity.

The shallow condition was contrasted against the *ignorance* condition, wherein participants read:

> Richard has been taught a great deal about astronomy and astrophysics. For example, if asked, he will tell you a lot of true things about the orbit of Mercury. He can accurately report Mercury's mass, volume, and average distance from the sun, and that Mercury's apparent orbit has some oddities.

> Richard can also tell you a lot of things about the theory of general relativity. General relativity explains some oddities in Mercury's observed orbit, but Richard was never told and never makes that connection, so he could not say anything about what explains those oddities. Everything Richard says about Mercury and general relativity is true, though. Nonetheless, he does not see any connections between the things he says about Mercury and about general relativity.

After reading one of the two vignettes[1], participants were randomly assigned to rate their agreement with one of the following two statements:

> Richard knows why Mercury has some oddities in its observed orbit

> Richard understands why Mercury has some oddities in its observed orbit

Answers were given on a seven-point Likert scale ranging from *Strongly Disagree* to *Strongly Agree* (coded as 1 and 7, respectively). To conclude the experiment, participants were asked five true-or-false reading comprehension questions to confirm that they had understood the vignette (e.g., "Richard has

---

[1] Some participants were assigned to a third *false* condition that we do not report here. In it Richard held a coherent set of false beliefs that he used to explain Mercury's observed orbit. On average, participants ascribed very little knowledge or understanding to Richard in this condition.

been taught some general relativity."), and were asked demographic and debriefing questions. Participants who made errors on the comprehension questions were excluded from further analysis.

Results

Responses were analyzed with a univariate ANOVA with belief condition (2: *ignorance, shallow*) and attribution (2: *knowledge, understanding*) as between-subjects factors (see Figure 1). This analysis found a main effect of belief condition, $F(1, 88) = 55.04$, $p < .001$, $\eta_p^2 = .39$, as well as a significant interaction between belief condition and attribution, $F(1, 88) = 6.85$, $p = .010$, $\eta_p^2 = .072$. Overall, participants attributed very little knowledge or understanding to Richard in the *ignorance* condition, and attributed significantly more in the *shallow* condition. However, knowledge attributions were significantly higher than understanding attributions in the shallow condition (post hoc Tukey HSD test: $p = .014$) and were not significantly different in the ignorance condition (post hoc Tukey HSD test: $p = .85$).
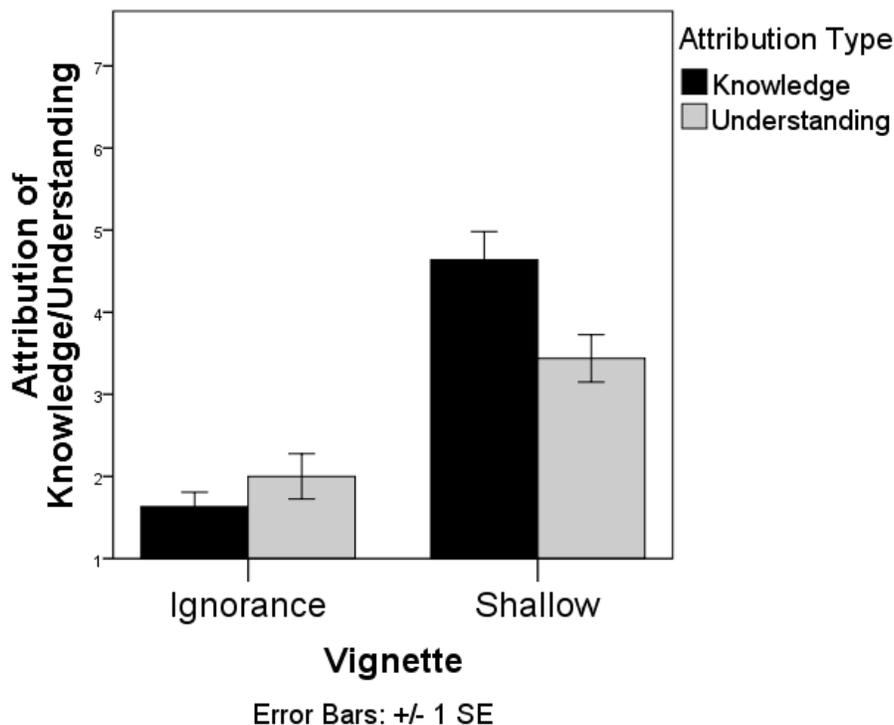
Fig. 1: Attributions of knowledge and understanding across conditions in Experiment 1. Error bars represent 1 SEM.

Discussion

Experiment 1 finds that participants are more inclined to attribute knowledge than understanding when the character has beliefs about an explanatory connection, but minimal explanatory depth. This is consistent with the prediction that attributions of understanding have higher demands than those of knowledge when it comes to explanatory depth.

Experiment 2

In this experiment, we aimed to replicate and extend the finding from Experiment 1 by considering different materials and more fine-grained variation in depth. In addition, depth was manipulated within subjects.

Methods

Participants Ninety-nine adults (39 female, mean age 32) were recruited through MTurk and participated in exchange for monetary compensation. An additional 43 participants were excluded prior to analysis following the same procedure used in Experiment 1.

Materials and Methods All participants read the same vignette, which described a student, "Edward," who progressively learned more about the impact event that caused the extinction of the dinosaurs. The vignette was presented in four stages. Participants were divided into two conditions, with half asked to make *knowledge* attributions after reading each stage of the vignette and half to make *understanding* attributions.

The initial stage of the vignette paralleled the *ignorance* condition from Experiment 1, but in a new domain:

> Approximately 65 million years ago, a large asteroid struck the Earth, spewing a large amount of ash and debris into the atmosphere. This debris prevented sunlight from reaching the surface of the Earth, which prevented plants from growing normally. This eventually led to the extinction of the dinosaurs.

> Edward is a student who has been taught about dinosaurs and about asteroids. But, as of Monday, he has never been taught about the connection between an asteroid impact and the extinction of the dinosaurs. If asked why the dinosaurs became extinct, Edward would not have been able to give a confident answer, and he would not mention an asteroid impact.

After reading this section, participants rated Edward's knowledge or understanding by reporting, on a 7-point Likert scale that ranged from "Strongly Disagree" to "Strongly Agree," their assessment of the statement:

> On Monday, Edward [knew/understood] why the dinosaurs became extinct

In the second stage, the following paragraph was added to the vignette, such that it then roughly matched the *shallow* condition from Experiment 1 in terms of depth:

> At a lecture on Tuesday, Edward learned that there is a connection between an asteroid impact and the extinction of the dinosaurs. If asked why the dinosaurs became extinct, Edward would say that an asteroid impact caused the dinosaurs to go extinct. If pressed for more detail, however, Edward could not elaborate and would not mention debris, blocked sunlight, or plants.

In the third stage, Edward learned more detail about how the impact led to extinction: It stirred up debris that blocked sunlight, hindered the growth of plant life, etc. Though Edward learned the *basic mechanism* by which the asteroid impact caused the extinction of the dinosaurs, the third vignette specified that Edward still lacked the ability to explain some points, namely why other asteroid impacts

had not caused extinction or why the debris blocked sunlight more disruptively than clouds do. Finally, in the fourth stage, Edward learned much more, such that he could explain the *complete mechanism* by which the asteroid caused the extinction of the dinosaurs (and, for example, why other asteroids did not cause the extinction of the dinosaurs).

After each stage participants again reported their agreement with statements of knowledge or understanding, with the day of the week changing appropriately.

Results

Responses were analyzed with a mixed ANOVA with attribution (2: *knowledge*, *understanding*) as a between-subjects factor and explanatory depth (4: *ignorance*, *shallow*, *basic mechanism*, *complete mechanism*) as a within-subjects factor (see Figure 2). This analysis revealed a main effect of explanatory depth, $F(1.94, 187.91) = 184.94$, $p < .001$, $\eta_p^2 = .66.$, as well as a significant interaction between depth and attribution, $F(1.94, 187.91) = 3.62$, $p = .030$, $\eta_p^2 = .036$.[2] Attributions increased monotonically with explanatory depth, starting near the bottom of the scale and ending near the top. Knowledge and understanding ratings did not differ significantly from each other in the ignorance condition, $t(97) = 1.20$, $p = .23$. However, in the three conditions involving greater explanatory depth, knowledge attributions were significantly higher than understanding attributions ($p \leq .050$).

Discussion

As expected, understanding and knowledge attributions both increased as a function of depth, but they did not do so in lockstep. At levels just above complete ignorance, participants were significantly more willing to attribute knowledge than understanding. This confirmed and generalized the results of Experiment 1, but still left the question of why this pattern exists—that is, why do knowledge and understanding attributions diverge in this way? In particular, do these attributions reflect distinct epistemic aims, and if so, what are they?

---

[2] As Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(5) = 83.73$, $p < .001$, the Greenhouse-Geisser correction for degrees of freedom was used, $\varepsilon = .65$.
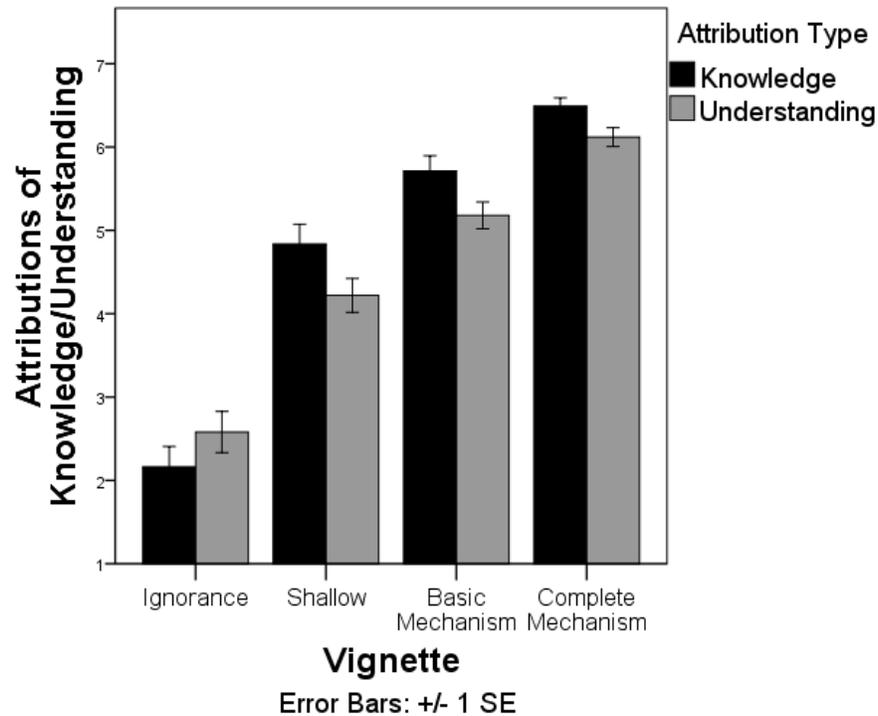
Fig. Attributions of knowledge and understanding at four levels of depth in Experiment 2.

Experiment 3

§5 Depth and Deference

Our contention is that understanding attributions are used to identify experts to whom one should defer, and that as a result we are differentially hesitant to ascribe understanding-why (relative to knowledge-why) when the beliefs constituting that understanding are not supported by access to deep explanatory information. One could, however, grant that the *Depth Difference* identifies the important distinction with regard to *Basis* (i.e. correctly predicts the factors that play into understanding attributions), but deny that we have correctly identified an answer to *Purpose* (that understanding-attributions are for picking out relevant experts).

To test this linkage we examined whether people's ratings of how much someone understood-why correlated with their ratings of how valuable that person would be to consult on a related problem. While the presence of such a correlation does not establish that identifying experts is what understanding-attributions are *for*, the existence of a systematic connection is certainly consistent with the proposal that the former might be guiding the latter. We will also test the stronger claim that

understanding attributions will track deference judgments *more closely* than attributions of knowledge do.

In order to demonstrate that the difference between knowledge-why and understanding-why attributions is general across domains, Experiment 3 involves stimuli from a different domain from Experiments 1 and 2. In addition, one might be concerned that, when participants are asked about knowledge and understanding in isolation, they have no frame of reference for their evaluations. We thus use a within-subjects design, in which the same participants are asked about both knowledge and understanding.

In sum, Experiment 3 aims to replicate and expand upon the findings from the previous two studies, as well as to directly measure whether understanding attributions are correlated with participants' willingness to defer to the understander regarding issues from the same domain. We did this by asking whether the character in the vignette would be a good person to consult about a related problem. In order to differentiate such selective deference from a general sense that the character in the deeper conditions was just more intelligent or generally informed, we also asked whether the character would be a good person to consult on a wholly disconnected matter.

Methods

Participants Sixty-five adults (34 female, mean age 35) were recruited through MTurk and participated in exchange for monetary compensation. An additional 63 participants were excluded prior to analysis following the same procedure used in Experiments 1 and 2.

Materials and Methods All participants read the same eight vignettes. Each vignette described a person who sits down at a computer, accidentally presses some keys on the keyboard, and causes an "Ω" character to appear on the computer screen. As in Experiment 2, vignettes varied in the explanatory depth of the character's grasp of the cause of an event, in this case the Ω character appearing on the screen. Unlike Experiment 2, vignettes were presented in random order, rather than in increasing order of depth, and the protagonist of each vignette was a different person. The complete text of the second of the eight levels of depth (analogous to the *shallow* condition from Experiments 1 and 2) read as follows:

> Dillon has learned almost nothing about computers. He has previously learned that pressing letter keys often makes the corresponding characters appear on the screen. He sits down in front of a computer and accidentally leans on the "Alt" and "Z" keys on the keyboard, causing an "Ω" character to appear in the document-editing program that is open.

> If asked why an "Ω" character appeared on the computer's screen, he would explain that leaning on the "Alt" and "Z" keys caused the "Ω" character to appear.

> If prompted for further detail, Dillon could not say more. He does not realize that a document-editing program was open and waiting for input and that an "Ω" character probably would not have appeared otherwise. He could not say anything about the code that makes up the document-editing program and how that code controls the computer, or about the various pieces of computer hardware involved and how they function, or about the low-level physics at work when a computer operates.

Vignettes for levels 3-8 increased causal depth, all the way to the "deepest" vignette (8), in which the character (Damien) was described as having "learned more than anyone else" about the relevant software and hardware, down to the low-level physics at work.

After each vignette, participants were asked to provide *both* knowledge and understanding attributions (with the order of presentation of these questions counterbalanced across participants, but constant for each participant), to answer a reading comprehension question (e.g., choosing whether the sentence "Dennis could give precise details about the CPU, like that it has separate logic and control units" is true or false), and to respond to three *deference* questions. These questions were designed to assess the perceived value of the protagonist as a source of information about the precise phenomenon at issue, about closely related processes, and about distantly related topics, respectively. Specifically, participants rated their agreement with each of the following statements (labeled in italics here, but not for participants):

> *Specific Fact Deference* [Name] would be a good resource for a person to consult if he or she were competing on a game show and had to accurately answer the question, 'How do you make an "Ω" character appear on a computer screen?'

> *Within-Domain Deference* [Name] would be a good resource for a person to consult if he or she needed help solving a technical problem with a computer.

> *Across-Domain Deference* [Name] would be a good resource for a person to consult if he or she needed help solving a difficult problem in economics.

After all eight vignettes, participants were asked to provide demographic and debriefing information.

Results

As in Experiment 2, we found that attributions for both knowledge and understanding increased with explanatory depth, and, consistent with Experiments 1 and 2, that participants were more willing to attribute knowledge than understanding at intermediate levels of depth (see Figure 3). Specifically, paired-samples t-tests at each level of depth revealed significantly-higher attributions of knowledge than of understanding for the 2nd, 3rd, and 4th depth levels (ps < .05).

Experiment 3 was also designed to assess the relationships between understanding attributions, knowledge attributions, and different kinds of deference. For each participant, we calculated the correlation between that participant's knowledge and understanding attributions and each of the three deference judgments across the eight levels of explanatory depth (see Table 1). We then analyzed these correlations as the dependent variables in a repeated-measures ANOVA with attribution (2: *knowledge*, *understanding*) and deference judgment (2: *specific fact deference*, *within-domain deference*) as within-subjects variables. We did not include the across-domains deference question in this analysis, as the majority of participants gave the same response (most frequently at the scale midpoint) to that question at all levels of explanatory depth. Accordingly, the correlations between knowledge and understanding judgments and the across-domains deference question were undefined for the majority of participants.[3]

---

[3] Of 65 participants, the correlation between the across-domains deference question and knowledge and understanding attributions was undefined for 34 of them. Among those participants for whom the correlation was

This ANOVA revealed a main effect of deference judgment, $F(1, 64) = 70.97$, $p < .001$, $\eta_p^2 = .53$, with correlations lower for *within-domain deference* than for *specific fact deference*, and a marginally significant main effect of attribution $F(1, 64) = 3.74$, $p = .058$, $\eta_p^2 = .055$, with correlations with knowledge lower than correlations with understanding. Importantly, there was also an interaction between attribution and deference judgment, $F(1, 64) = 10.22$, $p = .002$, $\eta_p^2 = .138$. Knowledge and understanding attributions were equally correlated with the specific fact deference question, $t(64) = -0.055$, $p = .96$. However, for the within-domain deference question, responses were significantly more correlated with understanding attributions than with knowledge attributions, $t(64) = -3.14$, $p = .003$. This confirms the key prediction that understanding attributions are more closely tuned than knowledge attributions to the epistemic aim of tracking reliable experts in a given domain.

Table 1: Mean correlations of value judgments with knowledge and understanding attributions in Experiment 3.

	Specific Fact Deference (Game Show)	Within-Domain Deference (Computer)	Across-Domain deference (Economics)[a]
Knowledge	.865	.678	.534
Understanding	.864	.735	.534

[a] The correlations between knowledge and understanding judgments and the across-domain deference question were undefined for the majority of participants because their responses were uniform across depth levels.
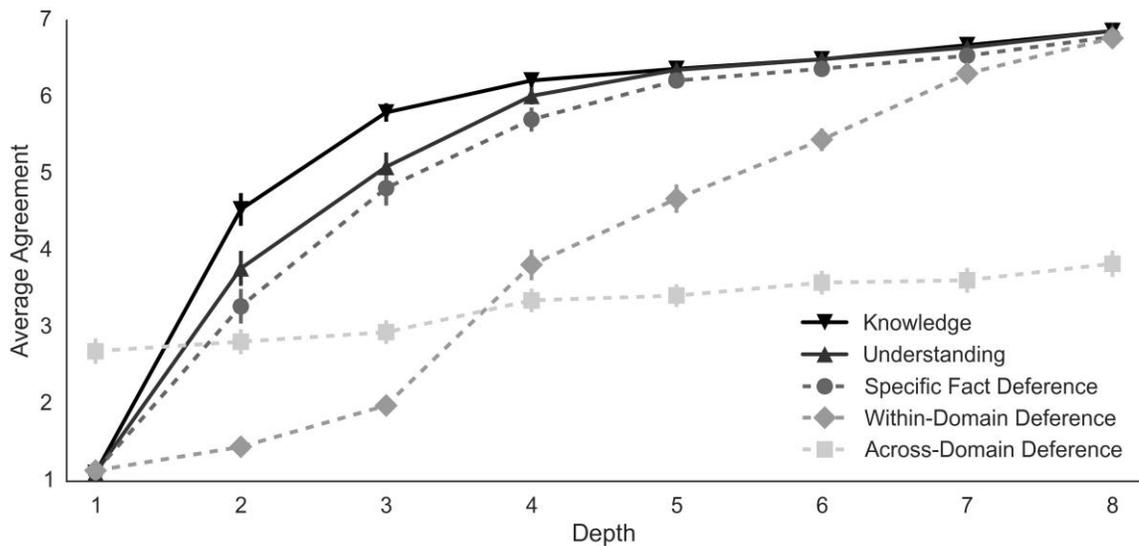

Fig. 3: Average attributions and deference judgments across eight levels of depth in Experiment 3.

---

defined, there was no significant difference between the correlation with knowledge attributions and the correlation with understanding attributions, $t(30) = 0.057$, $p = .96$.

Discussion

As predicted, attributions of understanding once again lagged attributions of knowledge for intermediate levels of explanatory depth. In addition, we found direct evidence that people attributed understanding, more so than knowledge, in proportion to their willingness to trust the person as a consultant on a related problem. The fact that this differential correlation was restricted to the within-domain deference question—and not found for the specific fact or between-domain deference questions—makes us more confident that the finding suggests a special relationship between understanding and deference within a domain of expertise, and not merely assessments of basic competence or general intelligence.

However, by asking about whom one should consult, Experiment 3 glossed over the question of whether the *basis* for deference is the expert's presumed personal skillfulness within the relevant domain or her "coaching ability." It is possible to be a valuable consultant on a problem either with or without possessing individual skill, and our question does nothing to disambiguate these possibilities.

Experiment 4

In Experiment 4, we were interested in replicating the results of Experiment 3, as well as disambiguating whether people were more likely to defer to understanders regarding questions in a related subject matter because they expected understanders to be more skillful themselves, or because they expected them to be valuable consultants for some other reason.

Methods

Participants Forty-seven adults (23 female, mean age 35) were recruited through MTurk and participated in exchange for monetary compensation. An additional 55 participants were excluded prior to analysis following the same procedure used in Experiments 1-3.

Materials and Methods Experiment 4 followed nearly the same procedure as Experiment 3. All participants read the same eight vignettes used in Experiment 3, and after each vignette they were asked to make both knowledge and understanding attributions (as well as to answer a reading comprehension question). However, the questions about the utility of the characters as experts were modified. First, we did not include the between-domain deference question, which frequently yielded uniform responses across vignettes. Second, the *specific fact* deference question, which was designed to measure the narrow ability to report that specific fact, was rephrased to more directly probe this ability. Third, the within-domain deference question, designed to measure *utility as an expert* in a particular domain, was similarly made more direct. Finally, two versions of each question were asked: one concerned with whether the character could *perform* a specific or general action and the other with whether the character could *advise* someone in performing a specific or general action. The four judgments were thus:

> *Specific Fact – Perform* If everything on the computer is working and set up to receive input, [name] would likely be able to make an "Ω" character appear on the screen.

> *Specific Fact – Advise* If everything on the computer is working and set up to receive input, [name] would likely be a good person to consult to find out how to make an "Ω" character appear on the screen.

Utility as Expert – Perform If there was a problem with the computer, [name] would likely be able to solve the problem.

Utility as Expert – Advise If there was a problem with the computer, [name] would likely be a good person to consult to find out how to solve the problem.

Finally, as in the previous three experiments, participants were asked to provide demographic information.

Results

First, we note that as in prior experiments, both knowledge and understanding attributions increased with depth (see Figure 4), with significantly higher knowledge attributions at intermediate levels: in this case, at levels 2, 3, and 5 ($ps \leq .05$).

Next, as in Experiment 3, we calculated the correlation between each participant's knowledge and understanding attributions and each of the subsequent judgments across the eight levels of explanatory depth (see Table 2). These were analyzed as the dependent variables in a repeated-measures ANOVA with attribution (2: *knowledge*, *understanding*), deference content (2: *specific fact, utility as expert*), and deference type (2: *perform, advise*) as within-subjects variables. This analysis yielded a significant main effect of attribution, $F(1, 46) = 17.88$, $p < .001$, $\eta_p^2 = .28$, with higher correlations for understanding ($r = .79$) than for knowledge ($r = .74$), a significant main effect of content, $F(1, 46) = 64.67$, $p < .001$, $\eta_p^2 = .58$, with higher correlations for the specific fact ($r = .86$) than for utility as an expert ($r = .67$), and two additional interactions. First, there was a significant interaction between attribution and deference content, $F(1, 46) = 8.28$, $p = .006$, $\eta_p^2 = .15$: Understanding was a better predictor than knowledge of utility as an expert ($r = .72$ versus $r = .62$), but not a better predictor for the particular fact ($r = .86$ versus $r = .85$). Second, there was a significant three-way interaction between attribution, content, and type, $F(1, 46) = 5.14$, $p = .028$, $\eta_p^2 = .10$. This interaction was likely driven by the perform/advise correlations for the specific fact, which were the only ones to exhibit a higher correlation for advising than performing.

Table 2: Mean correlations of value judgments with knowledge and understanding attributions in Experiment 4.

		Specific Fact	Utility as Expert
Knowledge	Perform	.868	.619
	Advise	.837	.619
Understanding	Perform	.853	.726
	Advise	.873	.716
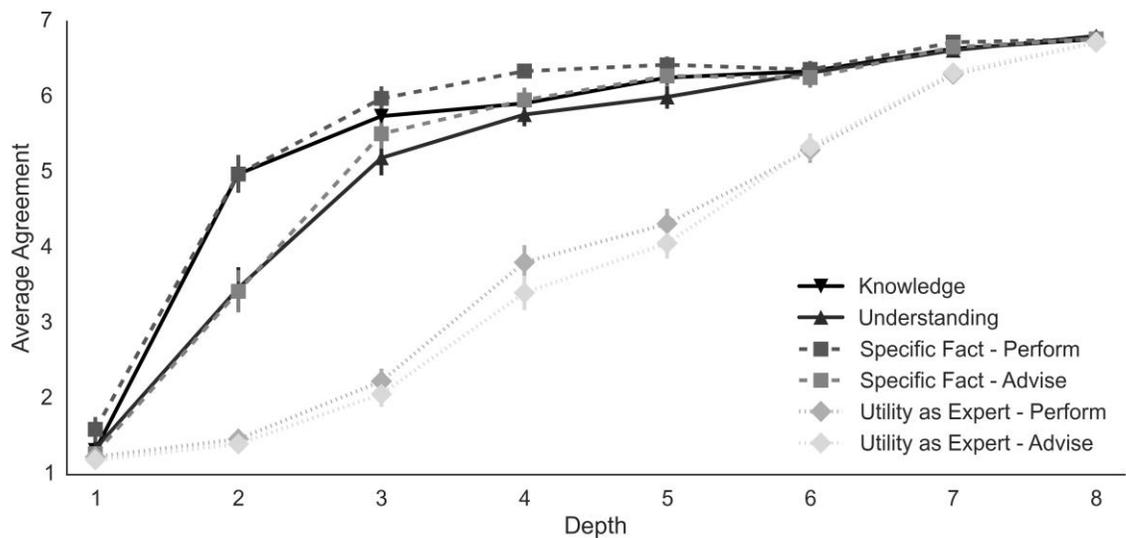
Fig, 4: Average attributions and deference judgments across eight levels of depth in Experiment 4

Discussion

In Experiment 4, participants continued to more closely tie understanding attributions with willingness to defer to the understander on a related problem. In our first time measuring, they also showed willingness to trust the understander to handle the problem herself, which suggests that the deference is reflective of a general respect for skillfulness in addition to respect for coaching ability.

General Discussion

Across four experiments, we found that people differentially withheld attributions of understanding (relative to knowledge) when the character possessed minimal explanatory depth. In Experiment 1, we found that people attributed significantly higher levels of knowledge than of understanding when the content of the relevant belief was a mere reference to there being some explanatory relationship. Experiment 2 extended this effect to a novel domain and to higher levels of depth. Experiment 3 revealed an important correlation: the extent to which people were willing to attribute understanding to a character was significantly correlated with the extent to which people would value that character as a useful consultant regarding a question within the same domain, and this relationship was significantly stronger for attributions of understanding than of knowledge. It was also restricted to deference *within* the relevant domain. This left open the question of whether people afforded such deference on the basis of understanders' presumed *skillfulness* or for some other reason. This was explored in Experiment 4, in which participants' assessments of understanding also significantly correlated with their willingness to trust a character to handle a problem directly.

These experiments shed light on the contours of our concept of understanding and point towards its functional role in cognition. In many ways, understanding-why seems to be closely related to knowledge-why, but our studies demonstrate that understanding-why carries with it a demand for greater explanatory depth. We further found that this differentially correlated with people's willingness to defer to understanders' expertise, at least in part because of an expectation of greater skillfulness. This suggests that people utilize understanding attributions to demarcate skillful experts to whom they

should defer. The presence of this depth-based asymmetry between understanding-why and knowledge-why is consistent with Pritchard's (2009) proposal, at least as regards our folk concepts, but challenges the prediction of Brogaard (2005). Our findings also provide a potential constraint on more neutral accounts of understanding that closely link understanding and knowledge, but without explicitly addressing questions of explanatory depth one way or the other (e.g., Grimm, 2006).

The present research continues a tradition from the literature in psychology of placing our willingness to consult others as a central feature of our epistemic lives (e.g., Keil 2003, Hofer & Pintrich 1997). However, our findings go beyond prior work by suggesting a particular mechanism – namely, the attribution of understanding – by which we track who has relevant expertise. Consistent with findings from Keil et al. (2008), our findings also suggest that patterns of deference are sensitive to the boundaries of disciplines as traditionally defined: we found differential correlations between deference and attributions of understanding (versus knowledge) within a domain, but not across domains (from computers to economics). Given the rich tradition of exploring folk epistemological concepts across development, our findings raise questions about how 'understanding' comes to possess its unique basis and purpose.

Of course, the present work does have several limitations. First, while we showed that people withhold attributions of understanding when a character's explanatory depth was moderate or shallow, this is far from establishing that such depth is *the* basis on which we attribute understanding. We attempted to rule out some confounds—such as general intelligence with the across-domains deference question—but of course our efforts were not exhaustive. Second, evidence of a correlation between understanding assessments and deference judgments does not show that the former have the function of tracking the latter. Future work should investigate additional bases for attributing understanding and knowledge, as well as other functions that might correlate with these attributions. It would also be valuable to identify the (potentially unique) epistemic aim of knowledge attributions – our focus here has been in characterizing understanding (as distinct from knowledge), but we have not offered an analysis of how attributions of knowledge function in our epistemic lives. Third, our participants were restricted to the folk; perhaps the judgment of expert epistemologists would look different, and give us better guidance to the true contours of understanding. Nonetheless, it's striking that our "naive" participants showed such systematic patterns of attribution, suggesting that while our conclusions may not capture those of expert epistemologists, they do reflect an intuitive epistemology that can be fruitfully investigated with the methods we employed. Finally, it will be important to relate these vignette-based attributions to actual "epistemic" behavior, including how attributions are used in communication, how they influence memory, and how they guide deference itself.

§6 Conclusion

For years—both anciently and quite recently—philosophers have been interested in what have come to be known as epistemological *Meno* problems. The problem was explaining why some particular state of grasping the world—traditionally knowledge, but also sometimes understanding—was valuable relative to the salient alternatives. In this paper we have defended one important dimension along which people use understanding-why and knowledge-why attributions differently, and have examined what possible purpose such a difference might serve. People expect understanders to have more explanatory depth than mere knowers, and people treat understanders as more appropriate experts to

consult on and solve difficult problems within the relevant domain. We suggest that this purpose is served by attributing understanding on the basis of explanatory depth.

References

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind.* Oxford university press.

Friedman, Michael (1974). Explanation and scientific understanding. *Journal of Philosophy 71*, 5-19.

Grimm, S. R. (2006). Is understanding a species of knowledge? *British Journal for the Philosophy of Science, 57*(3), 515-535.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science* New York: Free Press.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*, 88-140.

Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences*, *7*(8), 368-373.

Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. *Cognitive Science, 32*, 259-300.

Kelp, C. (2015). Understanding phenomena. *Synthese,* 1-18.

Kim, J. (1994). Explanatory knowledge and metaphysical dependence. *Philosophical Issues, 5*(Truth and Rationality), 51-69.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261-1277.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*, 309-328.

Kuhn, D., & Park, S. (2005). Epistemological understanding and the development of intellectual values. *International Journal of Educational Research, 43*, 111-124.

Machamer, P. K., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science* 67, 1-25.

Pritchard, D. (2009). Knowledge, understanding and epistemic value. *Royal Institute of Philosophy Supplement, 64*, 19-43.

Rosenberg, J. (1981). On understanding the difficulty in understanding understanding. In Herman Parret & Jacques Bouveresse (eds.), *Meaning and Understanding*. Berlin: W. De Gruyter.

Scriven, M. (1962). Explanation, prediction, and laws. *Minnesota Studies in the Philosophy of Science, 3*, 170-230.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition, 14*, 301-321.

Sosa, E. (2000). For the love of truth. In Linda Zagzebski (Ed.), *Virtue epistemology: Essays on epistemic virtue and responsibility* (pp. 49-62) Oxford University Press.

Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition, 124*, 272-283.

Strevens, Michael (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.

Weinberg, J. M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*, 429-460.

Weiskopf, D. A. (2011). Models and mechanisms in psychological explanation. *Synthese, 183*, 313-338.

Wilkenfeld, D. A. (2013). Understanding as representation manipulability. *Synthese, 190*, 997-1016.

Wilkenfeld, D. A. (2014). Functional explaining: A new approach to the philosophy of explanation. *Synthese, 191*, 3367-3391.