Invariantism, contextualism, and the explanatory power of knowledge

By Neil Mehta. Forthcoming in *Noûs*. Please cite the final published version.

It is often said that epistemology is the study of knowledge. One might naturally wonder why this is particularly worth studying: what is so special about knowledge? In recent decades, a number of philosophers have defended this answer to the question:

> The *Epistemic Theory of Mind*: Knowledge is part of the best overall framework for explaining behavior at the psychological level.

This would certainly give knowledge a special significance.

I find the Epistemic Theory of Mind to be very plausible; anyway, I will just assume for the sake of discussion that it is correct. What surprises me is that the framework is almost always conjoined with an invariantist theory of "know." I will argue that if we accept the framework at all, we have a very compelling reason to accept contextualism instead.

1. The Epistemic Theory of Mind, invariantism, and contextualism

Behavior can be explained at many levels: at the level of physics, chemistry, biology, and so on. Our discussion will focus on distinctively *psychological* explanations of behavior, such as explanations that appeal to beliefs and desires. The question is not what best explains any particular behavior in a one-off way. The question is what is the best *overall explanatory framework*. I will refer to any distinctively psychological explanatory framework as a *theory of mind*.

When we construct our theory of mind, we face important trade-offs. We can see the point by considering a series of increasingly stronger properties – say, belief, true belief, justified true belief, knowledge, and knowledge with certainty. This series is ordered by strength: having any property asymmetrically entails having all of the properties earlier in the series. Which properties in the series belong in our best theory of mind – our best overall framework for explaining behavior at the psychological level? Generalizations that appeal to stronger properties will be more powerful in one respect: they will yield a higher probability of the target behaviors given the presence of the psychological property. However, such generalizations will also apply in fewer cases. At the limit, we might explain every instance of behavior in terms of a maximally specific psychological property, a property that gets us as close to probability 1 of the behavior as the level of psychological explanation permits. Each such generalization would apply in a vanishingly small number of cases, with the result that we would need an

enormous number of generalizations, and an extraordinarily complex framework, to do the explanatory work that needs doing.

We find the reverse pattern of trade-offs if we limit ourselves to generalizations that quantify over weaker properties. As we weaken the psychological property required, we will get generalizations that apply more and more broadly. But we will simultaneously lower the probability of the behavior given that property (or reduce the specificity of the behavior that we predict). So perhaps we can just add some generalizations of both kinds to our theory? But, of course, every new generalization will increase our theory's complexity. It is a hard question which framework strikes the best balance among these desiderata.

It is widely agreed that the best overall framework will at least include generalizations about *beliefs* and *desires*. What else might figure in the framework, though? In recent decades, this answer has become increasingly popular:

> The *Epistemic Theory of Mind:* Knowledge is part of the best overall framework for explaining behavior at the psychological level.

Notice that the Epistemic Theory of Mind makes the most sense when linked with the idea that skepticism is false. The idea behind the theory is not that knowledge *would* have special explanatory power, if only people had it; it is that knowledge *actually* has such explanatory power. For that to be the case, ordinary subjects must actually stand in this epistemic relation, and with some regularity at that.

It is important to consider how the Epistemic Theory of Mind relates to a thesis that has been influentially endorsed by Timothy Williamson: the thesis that knowledge is a mental state. What is a mental state, exactly? Williamson says little by way of explanation except that "[p]aradigmatic mental states include love, hate, pleasure, and pain," as well as "attitudes to propositions: believing that something is so, conceiving that it is so, hoping or fearing that it is so, wondering whether it is so, [and] intending or desiring it to be so" (2000, p. 21). This does not give us a lot to go on. Still, it is at least clear that mental states are meant to be distinctively *psychological*, rather than merely biological, chemical, molecular, etc.

Jennifer Nagel elaborates helpfully on Williamson's thought by adding that "a capacity to register the mental states of others … enables us to explain and predict what [others] will do" (2013, pp. 276-277). Carlotta Pavese (2021, end of §2.1) is clearly working with a similar idea, and Adam Bricker simply piggybacks on the usage of Williamson and Nagel without further comment (2021, p. 2). Putting all of this together, the claim that knowledge is a mental state is at least the claim that knowledge is part of *some* kind of overall framework for explaining behavior at the psychological level. Which framework, though? Presumably what is most worth talking about is the *best* overall explanatory framework. Surely these theorists would be disappointed if knowledge were part of some pretty good overall explanatory frameworks, while being entirely absent from the very best one.

For these reasons, I will assume that knowledge is a mental state – as claimed by Williamson, Kornblith, Nagel, Pavese, and Bricker – just in case the Epistemic Theory of Mind is true. More generally, I will assume that:

A *mental state* is nothing more than a state that:

(i) is part of the best overall explanatory framework for explaining behavior, and
(ii) is psychological rather than merely biological, chemical, etc.

This is a claim about the *metaphysical nature* of mental states, not a claim about the *concept* of a mental state. Condition (i) is what ensures that the state in question is real – that it is natural, structural, or joint-carving, rather than gerrymandered or metaphysically arbitrary.[1] Condition (ii) then ensures that the state is distinctively *mental* rather than biological, chemical, etc.

This approach has a further virtue: it gives plausible truth conditions for Williamson's claim that believing truly is not a mental state (2000, pp. 27-30). For if believing truly is part of the best overall explanatory framework for explaining behavior, as per (i), then surely it explains behavior at the psychological level, not at the level of biology or chemistry, as per (ii). Thus, on the present approach, believing truly is not a mental state just in case it is not part of the best overall explanatory framework for explaining behavior. This seems exactly right: think of how strange it would be for Williamson to say, "Yes, believing truly is part of the best overall explanatory framework for explaining behavior at the psychological level. Nevertheless, believing truly is not a mental state."[2]

Still, it is easy to lose track of what is meant by the claim that knowledge is a mental state. Thus I will usually just speak of the Epistemic Theory of Mind itself.

My purpose in this paper is to consider how the Epistemic Theory of Mind interacts with another important epistemological debate, that between invariantists and contextualists. *Invariantists* think that "know" refers to a single, invariant relation no matter the context. *Contextualists* disagree; they think that "know" is context-sensitive, referring to different relations in different contexts. (At least, this is true when we restrict our attention to knowledge of propositions – knowledge-that – rather than knowledge of things. This entire paper operates under such a restriction.)

This gives us two very different ways of understanding the Epistemic Theory of Mind:

The *Invariantist Theory of Mind:* The word "know" refers invariantly to a single relation, knowledge. This relation is part of the best overall framework for explaining behavior at the psychological level.

The *Contextualist Theory of Mind:* The word "know" can, depending on the context, refer to countless different knowledge-like relations. Many or all

---

[1] On the notions of naturalness and structure, see Lewis (1983) and Sider (2011), respectively.
[2] I thank an anonymous referee for a helpful discussion of this point.

of these relations are part of the best overall framework for explaining behavior at the psychological level.

It is noteworthy that none of the theorists who accept the Epistemic Theory of Mind have committed to contextualism. By contrast, many of them, including Kornblith (2000), Williamson (2005), and Nagel (2010), have expressed at least strong sympathy for invariantism.[3]

I think that this is a mistake. My thesis is that if we accept the Epistemic Theory of Mind, then we have a powerful reason to prefer contextualism to invariantism.

2. From the Epistemic Theory of Mind to contextualism: an overview

In this section, I paint my argument for this thesis in broad strokes. I leave the fine brushwork for later sections.

An analogy will help us lock onto the basic idea of the argument. Suppose that being tall tends to make a person more attractive. How can we best incorporate this idea into a general explanatory framework? One option is to look for the single best place – 6.041 feet tall, let us suppose – to draw a line between those who are tall and those who are not. We would then say that those at or over the line tend to be more attractive than those under it. On this approach, our explanatory framework would of just one *binary generalization*:

> The *binary theory of height*: Tall people, i.e., people who are at least 6.041 feet tall, tend to be more attractive than everyone else.

This approach has substantial explanatory power.

But we might take a very different approach. Rather than making a single, binary division, we might appeal to the following *gradable generalization*:

> The *gradable theory of height*: The taller that someone is, the more attractive that person will tend to be.

The gradable theory is *much* more powerful than its binary counterpart. The binary theory can only explain differences in attractiveness between individuals whose heights are on opposing sides of the 6.041-foot mark. The gradable theory not only explains this, it also explains differences among individuals who are all below the line, or all above it. For example, it explains why being 6'1" makes someone more attractive than being 6'0", which makes someone more attractive than being 5'11", and so on. Consequently, it explains why the difference in attractiveness between someone who is 6'3" and someone who is 5'8" is *greater*

---

[3] What is more, in personal correspondence Nagel confirms that she is a "die-hard" invariantist, and Kornblith tells me that he is "very sympathetic to invariantism."

than the difference in attractiveness between someone who is 6'1" and someone who is 5'8". The binary theory is powerless to explain these facts.

There is a minor complication: past a certain point, becoming taller will tend to make one *less* attractive, not more. Thus, the generalizations under consideration need a ceiling: they should be applied only up to a certain height. (We could also add a floor, but I do not see the need: among adults, being shorter does seem to just continually decrease one's attractiveness.) This does not favor either the binary theory or the gradable one, since both theories will need this ceiling.

Which theory is better? To my mind, the gradable theory is the clear winner. Perhaps it is slightly more complex than its binary counterpart – although it is unclear to me whether appealing to a taller-than relation is more complex than appealing to a single invariant property of tallness. Still, the gains in explanatory power more than offset any potential loss in simplicity. There are just *so many* actual cases in which height differences correspond to differences in attractiveness, even though all of the individuals in question are on the same side of the 6.041-ft. line.

Notably, the property to be explained is gradable: a subject can be more or less attractive. This suggests a general lesson: often, gradable properties will be more aptly explained by gradable generalizations than by binary ones.

Bearing this in mind, return to the Epistemic Theory of Mind. As we will soon see, the advocates of this framework disagree about exactly which properties are distinctively explained by knowledge – but they all agree that these properties, whatever they are, are gradable. If we combine the Epistemic Theory of Mind with invariantism, then it is natural to explain these gradable properties by using *binary* generalizations, e.g., "Subjects who know *p* will tend to be significantly more persistent than subjects who do not know *p*." If we conjoin the Epistemic Theory of Mind with contextualism, we can instead appeal to *gradable* generalizations, e.g., "Standing in more and more stringent knowledge-like relations to *p* will tend to make the subject more and more persistent." These gradable generalizations perhaps involve a small loss in simplicity, but they make up for it with huge gains in explanatory power. That makes the gradable approach clearly superior in the end.

This is the *argument from gradability*. To express it more formally:

AG1. If the Epistemic Theory of Mind is correct, then what knowledge distinctively explains are non-maximal degrees of gradable properties.

AG2. If what knowledge distinctively explains are non-maximal degrees of gradable properties, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

AG3. If the Epistemic Theory of Mind is correct, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

In other words, given the Epistemic Theory of Mind, there is a strong reason to prefer contextualism to invariantism.

Let us inspect this argument more closely.

3. AG1 defended

Begin with:

> AG1.   If the Epistemic Theory of Mind is correct, then what knowledge distinctively explains are non-maximal degrees of gradable properties.

In support of AG1, I will review no less than five recent arguments for the Epistemic Theory of Mind. I will show that *all* of them posit, as explanatory targets of knowledge, gradable properties that are instantiated to a less-than-maximal degree.

First argument for the Epistemic Theory of Mind. This argument comes from Timothy Williamson. A central piece of the argument is this:

> A burglar spends all night ransacking a house, risking discovery by staying so long. We ask what features of the situation when he entered the house led to that result. A reasonable answer is that he knew that there was a diamond in the house. To say just that he believed truly that there was a diamond in the house would be to give a worse explanation, one whose explanans and explanandum are less closely connected. For one possibility consistent with the new explanans is that the burglar entered the house with a true belief that there was a diamond in it derived from false premises. For example, his only reason for believing that there was a diamond in the house might have been that someone told him that there was a diamond under the bed, when in fact the only diamond was in a drawer. He would then very likely have given up his true belief that there was a diamond in the house on discovering the falsity of his belief that there was a diamond under the bed, and abandoned the search. In contrast, if he *knew* that there was a diamond in the house, his knowledge was not essentially based on a false premise. Given suitable background conditions, the probability of his ransacking the house all night, conditional on his having entered it believing truly but not knowing that there was a diamond in it, will be lower than the probability of his ransacking it all night, conditional on his having entered it knowing that there was a diamond in it…. (2002, p. 62)

Williamson's argument proceeds via a direct comparison: he aims to show that, in a particular but representative example, behavior is better explained by knowledge than by true belief. In the surrounding passages, he uses similar examples to argue that knowledge also has more explanatory power than belief, true belief without reliance on false lemmas, stubborn belief, subjectively certain belief, and belief based on the best possible evidence. He takes this to constitute a strong initial case for the Epistemic Theory of Mind.[4]

What are the properties that Williamson takes to be distinctively explained by knowledge? I see up to three candidates. The first is *persistence*: Williamson's burglar persists in a particular action, continuing to search, because

---

[4] For the complete argument, see Williamson (2000, pp. 60-64).

he knows that the diamond is in the house. Persistence is obviously a gradable property, and what knowledge explains is not the maximum possible degree of it, but only some middling degree. For let us suppose that in our main case, the burglar never finds the diamond, it being hidden too well; he gives up at the break of dawn. There exists a vast continuum of other possible times when he could have given up: a minute past dawn, an hour earlier, at midnight, at noon.

A second potential explanatory target of knowledge is *risk-tolerance*. (It is not clear to me that Williamson has this property in mind, but let us consider it for the sake of thoroughness.) By remaining in the house, the burglar risks discovery and the attendant legal penalties. Why? Because he *knows* that the diamond is in the house. Risk-tolerance is another gradable property. The burglar could have been willing to risk less, or more, in an array of smooth possibilities. Perhaps the expected sentence for grand larceny is 10 years in prison; but this expected sentence could have been shorter or longer by days, months, years, or decades. And, again, knowledge will explain the property of risk-tolerance only to some non-maximal degree. For all his knowledge, the burglar need not be willing to stick around *on penalty of eternal torture.*

There is a third explanatory target of knowledge in Williamson's example: immunity to misleading counter-evidence. As the burglar continues his fruitless search, he gains more and more misleading evidence that the diamond is not in the house after all. Still, he retains his belief that the diamond is in the house. Why? Because he *knows* it is there. This explanatory target, too, comes in smooth degrees. A subject can be willing to dismiss just the tiniest amount of misleading counter-evidence, a moderate amount, or a lot. The burglar's knowledge surely does not immunize him to any amount of misleading counter-evidence whatsoever. It immunizes him to some particular, non-maximal amount of it.

Second and third arguments. Jennifer Nagel, drawing on extensive research in developmental psychology, argues that human beings *actually* routinely explain behavior in broadly Williamsonian ways; that is what our theory of mind is like. This theory has been very successful, she suggests, which provides substantial evidence that it is the best theory overall. In a similar vein, Adam Bricker argues that the brain attributes knowledge to others as if knowledge were a mental state rather than a composite of mental and non-mental conditions.[5]

We do not need to examine their arguments more closely. All that is important for our purposes is this. Williamson offers a particular account of *what kinds of commonsense psychological explanations* knowledge figures in, and the arguments of Nagel and Bricker are intended to show that commonsense explanations of these kind are, indeed, good ones. Thus, if Williamson says that the distinctive explanatory targets of knowledge are non-maximal gradable properties, then there is strong pressure for Nagel and Bricker to say so, too.

Fourth argument. Hilary Kornblith, another famous defender of the Epistemic Theory of Mind, has in mind an explanatory target that is quite different from those cited by Williamson. Kornblith's basic thought is that

---

[5] See Nagel (2013, especially §3–§4); Bricker (2021, especially §4).

knowledge plays an important role in explaining facts that are not about individual subjects, but about entire species. As he puts it:

> When we turn to an explanation of the cognitive capacities of the species, however, the theoretical enterprise we are now engaged in requires more than mere belief. We are no longer interested in explaining why a particular plover moved from its nest in a way that was bound to bring the predator's attention; instead we are interested in an explanation of how it is that members of the species are endowed with a cognitive capacity that allows them successfully to negotiate their environment. It is the focus on this adaptation of these cognitive capacities to the environment that forces us to explain the possibility of successful behavior, and it is the explanation of successful behavior that requires the notion of knowledge rather than mere belief. Knowledge explains the possibility of successful behavior in an environment, which in turn explains fitness. (2002, p. 57)

He adds:

> Notice that these explanations require more than just the category of true belief. If we are to explain why it is that plovers are able to protect their nests, we must appeal to a capacity to recognize features of the environment, and thus the true beliefs that particular plovers acquire will be the product of a stable capacity for the production of true beliefs. The resulting true beliefs are not merely accidentally true; they are produced by a cognitive capacity that is attuned to its environment. In a word, the beliefs are reliably produced. The concept of knowledge which is of interest here thus requires reliably produced true belief. (2002, p. 58)

In these passages, Kornblith is positing the following explanatory structure:

1. Plovers have regularly successfully negotiated their environments.

↑

2. Plovers have many true beliefs.

↑

3. Plovers have cognitive capacities that reliably produce true beliefs, i.e., cognitive capacities that generate knowledge.

Thus Kornblith thinks that it is the *reliability* of knowledge that gives it special explanatory power – in this case, the power to explain the general success of a species at negotiating its environment.

    I will pause to register a concern: I worry that Kornblith's theory entails that particular instances of knowledge *exist*, but it does not entail that they actually *explain* anything. To see the point, grant for the sake of argument that

knowledge is just reliably produced true belief. Then, yes, given *2* and *3* it will presumably be the case that plovers *have* many instances of knowledge. But what explanatory work are these instances of knowledge doing, exactly? *3* does not explain *2* by appealing to particular instances of knowledge, but by appealing to *cognitive capacities* that produce knowledge, which are quite different. And *2* does not mention knowledge at all. It mentions true beliefs *simpliciter*, without making any distinction between the knowledgeable ones and the ignorant ones. Thus, even taken on its own terms, Kornblith's theory only gives an explanatory role to *knowledge-producing cognitive capacities*. It leaves no explanatory role for *knowledge itself*. For this reason, I happen to prefer Williamson's argument for the Epistemic Theory of Mind.[6]

Still, I will bracket this concern below. All that we need to ask is this: are the explanatory targets of knowledge, according to Kornblith, gradable properties? Certainly. How many true beliefs plovers have, as per Kornblith's claim 1, and how regularly and successfully plovers have navigated their environments, as per Kornblith's claim 2, are both matters of degree. And is knowledge intended to explain something less than the maximal degree of these properties? Of course. As impressive as they might be, plovers do not have the maximum possible quantity of true beliefs, nor do they navigate their environments in the most successful possible manner.

Fifth argument. According to all of the arguments we have seen so far, knowledge figures importantly in *causal* explanations. Carlotta Pavese has argued that knowledge (also, or instead) figures importantly in *metaphysical* explanations. She begins by highlighting the importance of another psychological property, *agentive control*: this is necessary for skillful action and for intentional action, she thinks. She then argues that agentive control can be metaphysically analyzed partly in terms of practical knowledge – knowledge-how. Thus, on Pavese's view, the best framework will causally explain certain behaviors in terms of agentive control, which will in turn be metaphysically explained in terms of one form of knowledge. That is Pavese's reason for endorsing the Epistemic Theory of Mind.[7]

Does agentive control come in degrees? Of course. An expert pool player will have a very high degree of agentive control over her shots; a good pool player will have less; and I, sadly, have almost none. This should not come as a surprise. For Pavese thinks that agentive control is necessary for skillful action. But skill comes in degrees, and different degrees of skill will presumably require different degrees of agentive control. A non-gradable notion of agentive control could not do the explanatory work that Pavese has in mind.

Summary. In defense of the Epistemic Theory of Mind, it has been argued that knowledge can distinctively explain a wide variety of properties, from persistence, risk-tolerance, and immunity to misleading counter-evidence, to quantity of true beliefs and success in negotiating the environment, to agentive control. *All* of these properties are gradable, and, almost always, knowledge is meant to explain their presence to some less-than-maximal degree. That confirms

---

[6] I thank [name removed] for a helpful discussion on this point.

[7] See Pavese (2021, especially §3 and §7).

AG1: if the Epistemic Theory of Mind is correct, then the properties that are distinctively explained by knowledge are non-maximal degrees of gradable properties.

4. AG2 clarified

The remaining premise to consider is:

> AG2.   If what knowledge distinctively explains are non-maximal degrees of gradable properties, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

Before we can evaluate this premise, we should more carefully examine the two theories mentioned in the consequent of AG2.
The first theory. Recall:

> The *Invariantist Theory of Mind:* The word "know" refers invariantly to a single relation, knowledge. This relation is part of the best overall framework for explaining behavior at the psychological level.

The obvious way to develop the Invariantist Theory of Mind is to develop binary generalizations that appeal only to this single, invariant relation of knowledge. There is room to bicker over the precise form of these generalizations. Here are some reasonable candidates ("BG" stands for "binary generalization"):

> BG1.   Knowing p tends to make the subject persistent (or risk-tolerant, or …) to at least degree d when she acts on p.

> BG2.   Among subjects who act on p, those who know p tend to be much more persistent (or risk-tolerant, or …) than those who do not know p.

For now, it is not important exactly which binary generalization we pick. What *is* important is that this invariantist approach is not the only way of developing the Epistemic Theory of Mind.
A second theory. Another possible theory is this:

> The *Contextualist Theory of Mind:* The word "know" can, depending on the context, refer to countless different knowledge-like relations. Many or all of these relations are part of the best overall framework for explaining behavior at the psychological level.

The invariantist will deny the explanatory relevance of all but one of these knowledge-like relations. But even she should concede that these myriad relations *exist*: this much can be extracted from invariantist assumptions.

In particular, it is a staple of invariantist theorizing that knowing *p* requires having a belief with certain further gradable properties. For the sake of illustration, I will focus on the properties of safety, certainty, and reliability, but the reader is welcome to substitute other properties of this sort. The idea need not be that there are fixed degrees of safety, certainty, and reliability that are required for the subject to know any proposition in any circumstance. The idea need only be that *given a particular proposition and circumstance*, the subject's belief in that proposition must be safe, certain, and reliable to a certain minimum degree for the subject to know *that* proposition in *that* circumstance. Thus, if the single relation of knowledge exists, there must also exist a vast ocean of relations that are very similar to knowledge – for instance, each will require a non-accidentally true belief with some degree of justification – *except* that they require different minimum degrees of gradable properties such as safety, certainty, or reliability. I will refer to all such relations as *knowledge-like relations*.

We can partially order knowledge-like relations in terms of *stringency*. The intuitive idea is that a more stringent knowledge-like relation requires more safety, certainty, or reliability than does a less stringent one. To put things more precisely, knowledge-like relation *K+* is more *stringent* than *K-* when: (i) in every circumstance, having *K+* requires having at least as much of *each* relevant gradable property as does having *K-*, and (ii) in every circumstance, having *K+* requires more of *some* relevant gradable property than does having *K-*. In making this partial ordering, we are not assuming that these knowledge-like relations can be *analyzed* in terms of true belief plus these gradable properties. By analogy, we can partially order all shades of red in terms of saturation and brightness without assuming that any of the individual shades – crimson, scarlet, maroon – can be *analyzed* in terms of saturation and brightness. We could equally well analyze the particular degrees of saturation and brightness in terms of the shades that exhibit them, or deny that there is an analysis in either direction.

Once we have partially ordered knowledge-like relations by their stringency, we can construct a contextualist view on which the word "know" refers to different knowledge-like relations – sometimes more stringent ones, sometimes less stringent ones – in different contexts. In particular, we can endorse:

> *Referential Range*: For a very wide range of knowledge-like relations *K*, differing in stringency along the dimensions mentioned above, there is *some* metaphysically possible context *C* in which the English word "know" refers to *K*.

To get a better grip on how Referential Range works, contrast two contextualist theories of the word "tall." According to the first theory, "tall" will, in different contexts, refer to different degrees of a *single gradable property*, tallness. According to the second theory, "tall" will, in different contexts, refer to different *non-gradable* minimum height properties, like the property of being at least 5.671 inches tall. Referential Range is analogous to the second theory, not the first one. It posits, not a single *gradable* relation of knowledge, but countless *non-gradable* knowledge-like relations, as the different referents of the word "know" in different contexts. Gradability enters the picture elsewhere: different knowledge–like

relations require different degrees of gradable properties like safety, certainty, and reliability.

Referential Range is a perfectly reasonable commitment for the contextualist. Context-sensitive expressions often have a wide range of possible referents: think of how "large" can refer to tiny magnitudes ("large viruses") or enormous ones ("large galaxies"). In addition, Referential Range is compatible with a broad assortment of contextualist theories: it requires there to be stringency-based contextualist parameters for "know," but leaves wide open what further parameters there might be.

Referential Range suggests a new explanatory approach for the contextualist. Where the Invariantist Theory of Mind posits *binary* generalizations about *knowledge*, the Contextualist Theory of Mind can posit *gradable* generalizations about *knowledge-like relations*. For example (letting "GG" stand for "gradable generalization"):

> GG1. Standing in more and more stringent knowledge-like relations to p will tend to make the subject more and more persistent (or risk-tolerant, or …) in acting on p.[8]

At last, our two main Epistemic Theories of Mind are on the table. How do they compare to each other?

5. AG2 defended

My view is this:

> AG2. If what knowledge distinctively explains are non-maximal degrees of gradable properties, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

In this section, I will give my initial defense of AG2.

---

[8] Another relevant generalization is that standing in more and more stringent knowledge-like relations to p will tend to make the subject immune to more and more misleading counter-evidence. How should this generalization be expressed for the contextualist who thinks that $E = K$? The answer is that the contextualist's knowledge-like relations would then equally be evidence-like relations. Thus, the contextualist can posit the following two-part generalization:

> GG2. Standing in more and more stringent knowledge-like relations to p will tend to make the subject more and more immune to:
>
> (i) standing in more and more stringent evidence-like relations to counterevidence, and
> (ii) having stronger and stronger propositions as counterevidence.

Assume, for the sake of discussion, that knowledge does distinctively explain non-maximal degrees of gradable properties. For brevity's sake, I will focus solely on the properties identified by Williamson – persistence, risk-tolerance, and immunity to misleading counter-evidence – but the point generalizes to other gradable targets.

I grant that, given this assumption, the Invariantist Theory of Mind has one notable advantage: it is somewhat simpler than its contextualist competitor. In particular, its generalizations do not draw on many knowledge-like relations, but just one. I repeat that the invariantist should grant that the whole ocean of knowledge-like relations *exists*. The invariantist's advantage over the contextualist is not ontological. It is just that the invariantist has a more streamlined explanatory apparatus: only one relation, the unique relation of knowledge, figures in her explanatory generalizations. By comparison, the contextualist adds a bit more complexity.

But what explanatory power the contextualist gains in return! Just by taking these knowledge-like relations – whose existence is granted by all parties – and ordering them by stringency, the contextualist manages to explain a whole slew of differences that the invariantist cannot.

We can appreciate the point by considering all of the actual subjects who truly believe some well-known proposition – say, that climate change is primarily caused by human activity. These subjects will occupy a vast range of epistemic situations. Climatologists will have true beliefs that sail over the threshold for knowledge. Plenty of ordinary people will have true beliefs that amount to knowledge, but only by a small margin: think of those who know merely on the basis of catching some stray remarks from a teacher or news program. Others still will have true beliefs that fail, by margins small or large, to amount to knowledge. Perhaps these individuals will have heard varying amounts of counter-testimony by supposed experts.

The invariantist draws a single line between the knowledgeable and the ignorant. She can aptly explain many differences between the two. But she cannot explain the countless differences among those who know: she is silent about why the climatologist will tend to be much more risk-tolerant, persistent, and immune to misleading counter-evidence than the barely knowledgeable schoolchild. The invariantist is equally powerless to explain the countless differences among those who truly believe without knowing: for instance, why the person who has listened to only a little Fox News will tend to behave very differently from an enthusiastic viewer, even if both truly believe, on balance, that climate change is caused by human activity. The contextualist explains it all.

The contextualist can even explain certain differences between mere true belief and more epistemically robust true belief. The key is to think of true belief as simply the least stringent knowledge-like relation: it is the one that requires *no* safety, *no* reliability, and the lowest possible degree of certainty that is necessary for belief *simpliciter*. Thus, the contextualist correctly predicts that any knowledge-like relation more stringent than true belief will tend to make the subject more persistent, more risk-tolerant, etc.

How do these factors – simplicity versus explanatory power – weigh against each other, all things considered? To my mind, contextualism is the clear

winner: the invariantist leaves so much explanatory money on the table, and the contextualist spends so little to pick it all up. Just as it is worth investing in gradable generalizations about the explanatory power of tallness, so it is worth investing in gradable generalizations about the explanatory power of knowledge.

Against this, you might worry that the alleged virtue of the Contextualist Theory of Mind – its stronger and more fine-grained generalizations – is precisely its vice. Return to:

> GG1. Standing in more and more stringent knowledge-like relations to p will tend to make the subject more and more persistent (or risk-tolerant, or …) in acting on p.

This seems to entail that if we crank up *any* of the relevant gradable parameters – safety, certainty, reliability – without cranking down any of the others, the subject will become more persistent in acting on p. Is this bold claim really true?[9]

My reply is that GG1 does not entail the bold claim. For GG1 contains the phrase "tends to," and is therefore *much* weaker than the following claim:

> GG1*. Standing in more and more stringent knowledge-like relations to p *will* make the subject more and more persistent in acting on p.

To make vivid the difference between GG1 and GG1*, consider the graph below, which pertains to a particular hypothetical subject, in a particular situation, acting on a particular proposition p. The x-axis represents the stringency of the subject's knowledge-like relation to p; the y-axis represents her persistence in acting on p.
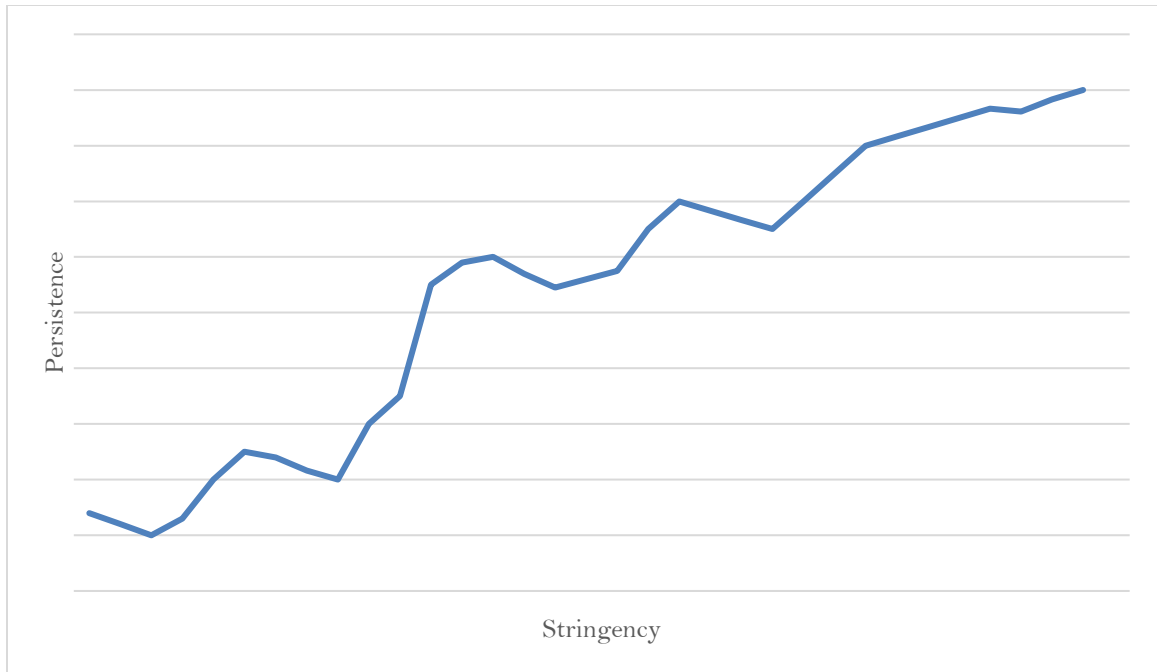
In the depicted situation, cranking up the stringency *sometimes* drops the subject's persistence, so GG1* is false. But GG1 is true: cranking up the stringency clearly *tends* to crank up the subject's persistence.

More generally, I submit that GG1 is not too bold; it is extremely plausible. Consider a burglar who just barely knows, by the invariantist's lights, that the diamond is in the house. This will make him *somewhat* persistent in looking for it, but not *maximally* persistent: at some point, if the diamond is well-enough hidden, he will give up. Now consider variants of the case in which the burglar has more and more evidence that the diamond is in the house: the housekeeper might be selling him information about what valuables are kept where; he might use audio surveillance equipment to hear the owners saying that they will just leave the diamond at home tonight; and so on. *Of course* this piling up of evidence will *tend* to make the burglar, if he is tolerably rational, more persistent in searching for the diamond.

There is theoretical room for doubt about these verdicts: it is not as though I have conducted double-blind controlled trials to verify them. Neither has Williamson, in giving his original version of the case. Still, my verdicts are extremely plausible, given what we know of human behavior, in the same way that Williamson's original verdict about the case is plausible, given what we know of human behavior.

On these grounds, I find AG2 very compelling. That, in turn, completes my initial defense of the argument from gradability: I tentatively conclude that, given the Epistemic Theory of Mind, there is a strong reason to prefer the Contextualist Theory of Mind to its invariantist counterpart.

But let us consider some potential objections.

6. Interlude

Here is the argument again:

> AG1. If the Epistemic Theory of Mind is correct, then what knowledge distinctively explains are non-maximal degrees of gradable properties.
>
> AG2. If what knowledge distinctively explains are non-maximal degrees of gradable properties, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.
>
> AG3. If the Epistemic Theory of Mind is correct, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

As far as I can tell, AG1 is watertight; it just expresses the commitments of those who defend the Epistemic Theory of Mind. The rest of the argument is more vulnerable.

Against AG2, two objections may be raised: first, that the knowledge-like relations that I have described do not explain behavior *at the psychological level*, and thus are not mental states; second, that I have underestimated the explanatory resources of invariantism. I will consider these objections in §7-§8, respectively.

In addition, some will object that AG3, even if it is true, is insignificant. The idea is that the very best theory of mind would appeal only to the safety, certainty, and reliability of the subject's *true beliefs*; knowledge, or knowledge-like relations, would be jettisoned entirely. I will engage with this concern in §9.

7. Objections to the idea that knowledge-like relations are mental states

Recall:

> AG2. If what knowledge distinctively explains are non-maximal degrees of gradable properties, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

The Contextualist Theory of Mind claims, among other things, that many different knowledge-like relations are part of the best overall framework for explaining behavior *at the psychological level*. In §1, I argued in turn that this is true just in case many different knowledge-like relations are mental states. Thus, one way to object to AG2 is to argue that knowledge-like relations fail to explain behavior *at the psychological level*, and thus are not mental states. In this section, I will first lay out and rebut several particular objections of this kind. Then I will give a principled argument against any objection of this kind.

First objection. As a warm-up to the objection, consider this argument from Williamson (2001, pp. 28-30; "M" stands for "mental"):

M1. If a concept C is the conjunction of the concepts $C_1 \ldots C_n$, then C is mental only if each of the concepts $C_1 \ldots C_n$ are mental.

M2. The concept of believing truly is the conjunction of the concepts of believing and being true.

M3. The concept of believing truly is mental only if each of the concepts of believing and being true are mental. (From M1, M2.)

M4. The concept of being true is not mental.

M5. The concept of believing truly is not mental. (From M3, M4.)

Now consider any knowledge-like relation, R, other than invariantist knowledge. It is natural to wonder whether we could argue, in similar fashion, that the concept of R is not mental.[10] The argument would begin with the same premise:

M1. If a concept C is the conjunction of the concepts $C_1 \ldots C_n$, then C is mental only if each of the concepts $C_1 \ldots C_n$ are mental.

The next premise would need to take the following form:

The concept of R is the conjunction of the concepts of knowledge and ____.

However, it is unclear how to fill in this blank. As I have defined things, R is a relation that is similar to knowledge, *except* that it requires different minimum degrees of various gradable properties, such as safety, certainty, or reliability. This definition is not a conjunction of conditions – it is not of the form, *R is a relation that is knowledge and __* – for R is not, by the invariantist's lights, a relation of knowledge. Thus there is no threat to the Contextualist Epistemic Theory of Mind here.

Second objection. The invariantist might press this objection ("MC" stands for "mental concept"):

MC1. A state S is mental just in case there can be a mental concept of S.[11]

MC2. There cannot be a mental concept of knowledge-like relations (other than invariantist knowledge).

MC3. Knowledge-like relations (other than invariantist knowledge) are not mental.[12]

Is this argument sound?

I grant MC1: I regard it as more or less analytic. (We might roughly define a *mental concept* as a concept of a mental state, regarded in its mental aspect.) But I cannot find any particular evidence for MC2. For example, you might argue that the concept of any particular knowledge-like relation (besides invariantist knowledge) is a cousin of the concept of knowledge, but many cousins of mental

concepts are not themselves mental concepts. I agree, but observe that *some* cousins of mental concepts *are* themselves mental concepts: the concepts of pain and pleasure are cousins, as are the concepts of pain and shooting pain. So this line of reasoning gives us little evidence either for or against MC2.

Thus the present objection is unsupported.

Stepping back. At this point, you might think that, even if we do not yet have evidence that knowledge-like relations are *not* mental states, neither do we have evidence that they *are* mental states. Thus it would seem unfair for me to assume this. Can I make a positive case that knowledge-like relations are mental states?

I reply that I have *already* made this case. For I have been arguing all along that:

> MS1. Knowledge-like relations are part of the best overall explanatory framework for explaining behavior.

("MS" stands for "mental state.") In addition, I defended the following claim in §1:

> MS2. For any x, if x is part of the best overall explanatory framework for explaining behavior, then x is a (genuine, natural, structural) state.

It follows that:

> MS3. Knowledge-like relations are (genuine, natural, structural) states.

And it is extremely plausible that:

> MS4. If knowledge-like relations are (genuine, natural, structural) states, then knowledge-like relations are mental states.

For assuming that knowledge-like relations are (genuine, natural, structural) states in the first place, it is hard to see what other kind of states they could be. They at least look like relations to propositions; they figure in explanations of action; and they are clearly not biological, chemical, or atomic states.

MS3 and MS4 entail:

> MS5. Knowledge-like relations are mental states.

This is our desired conclusion.

Notice that MS2 and MS4 cohere well with my claim in §1 that a *mental state* is nothing more than a state that (i) is part of the best overall explanatory framework for explaining behavior, and (ii) is psychological rather than merely biological, chemical, etc. The idea is that because knowledge-like relations satisfy (i), they are genuine *states*; and because they satisfy (ii), they are *mental* states rather than states of some other kind.

In addition, it is clear that gerrymandered relations to propositions, like *either believing that snow is white or having hairy arms,* are not mental states. Thus,

we had better not be able to construct an argument, parallel to the one above, for the conclusion that they *are* mental states. And, indeed, we cannot: the parallel of MS1 would be obviously false. More generally, gerrymandered relations to propositions are not (real, natural, structural) states at all; *a fortiori*, they are not mental states.

I conclude that there is strong evidence that knowledge-like relations are mental states, and no particular evidence to the contrary.

8. Objections that appeal to more sophisticated invariantist theories

Return once more to:

> AG2. If what knowledge distinctively explains are non-maximal degrees of gradable properties, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

While defending this premise, I considered only a few simple versions of the Invariantist Theory of Mind. However, the invariantist might attack AG2 by developing more sophisticated versions of the Invariantist Theory of Mind. In this section, I consider many variations of this strategy.

First option. Distinguish between two forms of invariantism. Consider a subject who truly believes that *p*. Does she know that *p*? According to *intellectualist invariantism*, the answer to this question depends *solely* on factors that are, in some way, truth-related – for instance, how much evidence the subject has, and how safe, certain, and reliable her true belief is. According to *subject-sensitive invariantism*, this is not the case; the answer to the question depends also on such factors as how much is at stake for the subject. Thus a single subject might cease to know that *p solely* because it has become more important to her whether *p* is true, without any change in her evidential position, her confidence in *p*, etc.[13] Grant for the moment that the contextualist's theory of mind is better than that of the *intellectualist* invariantist. You might still reasonably wonder: can the *subject-sensitive* invariantist do better?

Reply. I do not see how. All that intellectualist and subject-sensitive invariantists disagree about is the *precise location* of the line between knowledge and ignorance: about whether subjects can be on different sides of the line despite being identical in their evidence and in the safety, certainty, and reliability of their true beliefs. But my core concern has nothing to do with precisely where to place the line between knowledge and ignorance. My concern is that we want to explain gradients in behavior, and *no* single line can do that, no matter where you put it. To put things another way, the subject-sensitive invariantist agrees with her intellectualist counterpart in positing only binary generalizations; it is just that

---

[13] Intellectualist invariantism is the traditional view. For some detailed defenses of subject-sensitive invariantism, see Stanley (2005); Fantl and McGrath (2009).

her binary generalizations play out somewhat differently in particular cases. This does not address the problem, which is that we *do not need* binary generalizations, we need gradable ones.

Second option. It is common for invariantists to appeal, not just to knowledge, but also to iterations of it.[14] This might inspire the invariantist to reframe her generalizations. For instance, she might suggest that when acting on p, having $n + 1$ iterations of knowledge of p tends to make a subject more persistent, risk-tolerant, etc. than having n iterations of knowledge of p. The invariantist's original generalizations then drop out of the special case in which $n = 0$.

Reply. This approach suffers from a pair of problems.

One problem is that the Epistemic Theory of Mind is not just intended to apply to human beings. It is supposed to apply to non-human animals, too. But the current approach does not extend well to these cases. The trouble is that non-human animals can behave very differently depending on their epistemic positions, even though they lack the intellectual sophistication to conceive of themselves as knowing anything, and thus cannot have iterations of knowledge. Think of how a dog might tend to become more and more persistent in seeking an especially valued but difficult-to-access piece of food as we increase the quality of the dog's perceptual evidence that such food is nearby.

There is a second problem. Yes, the current approach lets the invariantist explain the differences among those who merely truly believe, those who know, and those who know that they know. But it compresses large differences within each category, while expanding tiny differences across categories. Compare (i) someone who truly believes p with no justification whatsoever and (ii) someone who truly believes p on the basis of substantial evidence while falling barely short of knowledge. These subjects will tend to behave very differently, but the current approach makes no distinction between them. Now compare the subject described in (ii) with (iii) someone who has just a bit more evidence, and thus just barely knows p. The current approach makes a large distinction between them, even though their behaviors will tend to differ only slightly. These are just the sorts of problems that arose for the original version of the Invariantist Theory of Mind. The root cause of the problem is that these approaches provide explanatory stairs when we need ramps.

The lesson is that we really do want gradable generalizations. This is something that the remaining invariantist options all provide.

Third option. The invariantist might suggest this pair of gradable generalizations. First, the closer a true belief is to knowledge, the more persistent, risk-tolerant, and so on the subject will tend to be. Second, the further "past the post" of knowledge a true belief gets, the more persistent, risk-tolerant, and so on the subject will tend to be. The invariantist might hope that this framework will enjoy the full explanatory power of its contextualist rival.

Reply. One small concern with this invariantist approach is that it is, in one respect, actually more complex than its contextualist counterpart. Wherever the

---

[14] See for example Williamson (2005, p. 231); Turri (2010, p. 87).

contextualist uses one generalization, the invariantist uses two: one about true beliefs that fall short of knowledge, another about true beliefs that go past the post of knowledge.

But there is a second problem, which goes much deeper: the line between knowledgeable and non-knowledgeable true beliefs has become explanatorily irrelevant. All of the explanatory work is being done by the partial ordering of knowledge-like relations by stringency. In other words, on this approach, the invariantist is no longer using the single relation of knowledge to do any significant work whatsoever. The result is that the current invariantist approach is not a way of developing the Epistemic Theory of Mind, but a way of abandoning it entirely.

Fourth option. It is sometimes held that a subject's evidence is identical to what she knows: $E = K$.[15] This idea is perfectly available to contextualists and invariantists alike, but the invariantist might think to leverage this idea into a new theory of mind.

The thought is this. The invariantist might grant that there are many differences among subjects who truly believe p. But she might explain these differences, not in terms of whether the subject knows p itself, but in terms of how much evidence the subject has in support of p. This might seem to let the invariantist introduce the kinds of gradable generalizations that are needed – for example, the generalization that the more evidence a subject has for p, the more persistent she will tend to be when acting on p. At the same time, this invariantist approach hangs onto the Epistemic Theory of Mind, since evidence is itself identified with (invariant) knowledge.

Reply. I regard this as a large improvement on the previous invariantist attempts. Still, I think that the contextualist approach is significantly better. First, notice that this invariantist approach is much closer in complexity to the contextualist approach: both include an array of gradable generalizations that are structurally identical. Invariantism thus loses part of its former advantage in parsimony. Still, the invariantist retains some edge: she appeals only to the single relation of knowledge, not to a whole array of knowledge-like relations.

But there is a much more important difference between the invariantist and contextualist approaches: only the contextualist can explain the many differences among subjects who have true *and non-inferential* beliefs in the same proposition. The problem for the invariantist is that true non-inferential beliefs are not based on further evidence, so her explanatory apparatus gets no traction in such cases.

Suppose for example that Mata Hari is fleeing from pursuers. She runs past a park bench, looking at it for a signal from her handler that will tell her where to go next. Just on the basis of vision, she forms the non-inferential true belief that the bench has a particular chalk symbol on it. From this, she infers, truly, that her handler is waiting for her with a running car on Elm Street. We can imagine many versions of this case that differ in Mata Hari's epistemic position with respect to the proposition that the bench has the particular chalk symbol on it. In some cases,

---

[15] See Williamson (2000, ch. 9).

she is running so fast that she can only catch a quick glimpse from the corner of her eye; she truly believes, but does not know, that the symbol is there. In other cases, she has gotten a slightly longer look and just barely knows. In still other cases she has gotten a pretty good look or an excellent one. There will be large differences among these different cases: the better Mata Hari's epistemic position with respect to the proposition, the more persistent, risk-tolerant, and so on she will tend to be.

The invariantist's idea was to explain how persistent (e.g.) Mata Hari will tend to be by appealing to how much evidence she has that her handler is waiting on Elm Street. But, in this case, Mata Hari has no further evidence for this proposition. She just non-inferentially believes it. Thus she either knows it, in which case it is, by my opponent's lights, evidence itself, or she does not know it, in which case she has no evidence. To be sure, this does let the invariantist explain why Mata Hari tends to be more persistent when she knows that the chalk symbol is there than when she does not know. But the invariantist exaggerates the differences between cases where Mata Hari barely fails to know and cases where she barely knows. What is more, the invariantist cannot at all explain the many differences in Mata Hari's persistence across cases in which she knows, as well as the many differences across cases in which she truly but non-knowledgeably believes. The contextualist pockets the money that the invariantist leaves on the table.

How much money is that, exactly? A lot. Perception gives any normal subject a constant stream of non-inferential true beliefs. These beliefs vary greatly in their behavioral effects, depending on the quality of the perception at issue: think of the difference between seeing something clearly and seeing it from the corner of your eye or from a great distance. These cases feature gradable variations in the subject's persistence, risk-tolerance, and immunity to misleading counter-evidence. The invariantist approach draws only binary distinctions, which are inadequate to the explanatory task. The contextualist approach gives us the gradable distinctions that we need.

Fifth option. The invariantist might propose gradable generalizations such as the following: the greater the *minimum likelihood* the subject knows p to have, the more persistent (risk-tolerant, …) she will tend to be when acting on p.

To appreciate the idea, suppose that Theodore has been murdered. The veteran detective Poirot investigates, and learns the following claims *in sequence*:

A1. The butler's footprints are on the scene.
A2. The butler was heard shouting at the victim.
A3. The bloody knife is in the butler's quarters.
A4. …

The invariantist might suggest that, *before even commencing his investigation*, Poirot knows the following conditionals:

B1. If the butler's footprints are on the scene, then it is at least mildly likely that the butler did it.

B2. If the butler's footprints are on the scene and the butler was heard shouting at the victim, then it is at least somewhat likely that the butler did it.

B3. If the butler's footprints are on the scene and the butler was heard shouting at the victim and the bloody knife is in the butler's quarters, then it is at least moderately likely that the butler did it.

B4. …

Thus, Poirot gradually becomes able to learn these increasingly informative propositions:

C1. It is at least mildly likely that the butler did it.

C2. It is at least somewhat likely that the butler did it.

C3. It is at least moderately likely that the butler did it.

C4. …

And this, at last, might explain why Poirot will tend to become increasingly persistent, risk-tolerant, and so on when acting on the proposition that the butler did it.

The invariantist might suggest that many cases are like this. *At the start of inquiry,* scientists will often know that a particular series of experimental results would monotonically raise the minimum likelihood of a certain scientific hypothesis; historians will often know that a particular series of historical records would monotonically raise the minimum likelihood that certain historical events occurred; and ordinary people will often know that a particular series of behaviors would monotonically raise the minimum likelihood that their friend was upset. Thus, as these subjects gradually accumulate these pieces of evidence, they come to know increasingly stronger truths about the minimum likelihood of the relevant proposition, p.

Note well: none of this requires the subject to know *what her evidence is.* A1 is not the proposition *that it is part of the subject's evidence* that the butler's footprints are on the scene; it is just the proposition that the butler's footprints *are* on the scene. Note also that, although the B-truths and the C-truths are about *epistemic* likelihood – likelihood given one's evidence – the subject can know these truths without knowing anything about epistemic likelihood. (Analogy: even though water *just is* H_2O, a subject can know truths about water without knowing anything about H_2O.)

Are these invariantist generalizations up to the task?

Reply. I think not: by themselves, these generalizations do not explain important differences among subjects who all know the same proposition. Consider once more the proposition that climate change is primarily caused by human activity. Compare (i) a subject who barely knows this on the basis of testimony to (ii) a climate scientist who fully appreciates the enormous weight of evidence for this proposition. Both of these subjects know the same proposition, full stop, with no probabilistic modifier attached. But they will tend to be radically

different in, for instance, their immunity to misleading counterevidence. The current approach leaves us powerless to explain these large differences.[16]

Sixth option. The invariantist's fourth option was to say that the more evidence – understood as invariant knowledge – the subject has in support of p, the more persistent she will tend to be. The problem was to explain differences among subjects who have no further evidence for p. The fifth option was to say that the higher the minimum likelihood that the subject knows p to have, the more persistent she will tend to be. This produced a different problem, the problem of explaining differences among subjects who all know the same proposition. A natural thought is to combine these generalizations, thereby solving both problems at once. The invariantist can explain differences among subjects who have no further evidence for p by saying that they differ in how likely, at a minimum, they know p to be. And she can explain the differences among subjects who all know the same proposition p by observing that some of these subjects can have much more evidence for p than others.

Reply. I regard this as the best version of the Invariantist Theory of Mind. Still, on reflection, I think that the contextualist approach proves superior. The basic reason is that the current approach draws lines that are profoundly arbitrary. These lines reflect the underlying terrain about as well as the state boundaries reflect the natural geographical breaking points in the United States: occasionally well, but mostly at random.

One symptom of the problem is that the invariantist approach "doubles up" its predictions in certain cases. Suppose that some proposition p is true. Compare two subjects who arrive at different pieces of knowledge concerning p via inference: the first subject inferentially knows that p is at least 60% likely to be true; the second subject, having some extra evidence, inferentially knows that p is at least 70% likely to be true. Both of the invariantist's generalizations apply to this case: the subjects differ in how likely they think p is, *and* in how much evidence they have in support of p. As things stand, then, the invariantist should predict a double boost, so to speak, in the persistence of the second subject over the first subject. That prediction is incorrect: having more evidence is what *causes* the second subject to know something more specific about the minimum likelihood of p. Thus, the boost from the knowledge about minimum likelihood should already soak up any boost from increased evidence.

The invariantist can patch over the problem: she can add an exception for cases in which one generalization applies *because* the other generalization applies.

---

[16] Return to:

> B1. If the butler's footprints are on the scene, then it is at least mildly likely that the butler did it.

An anonymous referee raises a further concern for my opponent: in some contexts, B1, far from being knowable, will be false. For the fact that the butler's footprints are on the scene might be only part of the subject's evidence. Given the subject's *total* evidence, p might not be even mildly likely.

But it is worrying that the invariantist needs to apply an *ad hoc* patch like this in the first place.

Another symptom of the problem has to do with differences among true beliefs that are *simultaneously* non-inferential and non-probabilistic. Consider a series of subjects who each see a patch of color for different periods of time – anything from 1 millisecond to 10 seconds, let us say. We may suppose that each subject comes to believe, just on the basis of perception, that the patch is red. These are all, we may stipulate, cases that the invariantist would understand as involving *full* – if sometimes irrational – belief. None of them are, by the invariantist's lights, cases of merely partial belief.

Neither of the invariantist's generalizations will make any gradable distinctions among these true beliefs, for they are all non-inferential *and* their contents are all non-probabilistic. The invariantist will only make a binary distinction between those true beliefs that are knowledge and those that are not. That is unfortunate. For these subjects can be placed along a spectrum by length of visual exposure to the patch of color, and there can easily be relatively smooth differences in how resistant their true beliefs are to misleading counter-evidence. These differences cry out to be explained by a gradable generalization.

Yet another symptom of the problem is that the invariantist draws a seemingly arbitrary line between probabilistic and non-probabilistic knowledge. Consider a subject who is steadily gaining evidence for *p*, the proposition that climate change is primarily caused by human activity. The evidence can accumulate in a very smooth, gradual way, but the invariantist draws a sharp line at the point where the subject switches from knowing that *p* is likely to knowing *p*, full stop. At this point, the invariantist switches from one explanatory generalization to another. Why? It is very hard to see why it is worth drawing a line here: the subject can continue to accumulate evidence in a smooth way long after she has achieved knowledge of *p*. There is no reason to think that anything "clicks into place" at the precise moment when the subject achieves knowledge. The line seems perfectly arbitrary. The contextualist can avoid drawing such a line by saying that *all* (human) knowledge is probabilistic: the subject just keeps bumping up her credence in *p*, never reaching credence 1.

The arbitrariness just described has to do with the invariantist's line between knowing *p* and not knowing *p*. There is one last symptom of the problem. It is that the invariantist draws an equally arbitrary line between knowing and not knowing the underlying propositions that constitute one's evidence. To see the point, suppose that there is a single evidential proposition, *e*, that provides knowledge-level support for *p*. Consider a series of cases in which we gradually strengthen the subject's epistemic position with respect to *e*: in every case the subject knows *e*, but in some cases she just barely knows it, while in other cases she sails past the threshold of knowledge. We may suppose that in all of these cases, the subject's knowledge of *e* is non-inferential. The invariantist will treat all such cases alike: the subject always has exactly the same evidence, *e*, for the same proposition, *p*, and her belief is always flat-out (by the invariantist's lights) rather than probabilistic. But there can be *enormous* differences in the subject's persistence, risk-tolerance, and immunity to misleading counter-evidence as we

strengthen the subject's grip on her evidence! The invariantist leaves these differences unexplained.

These are serious problems. Taken cumulatively, they give us strong reason to prefer the contextualist alternative.

A complication. Across this section, I have repeatedly assumed that a subject who knows p can continue to accumulate evidence, consisting of propositions $E_1 \ldots E_n$, for p. Some will object that this is impossible. The worry relies on the following principles:

> The *Evidence-Probability Link*: E is evidence for p just in case the epistemic probability of p given E is greater than the epistemic probability of p alone.

> $E = K$: One's evidence consists of all and only the propositions that one knows.

Given these principles, when a subject knows p, the epistemic probability of p is already 1; it cannot be raised by any further evidence. Thus, this subject cannot possibly gain *more* evidence for p.[17]

I grant that the Evidence-Probability Link characterizes one important sense of the phrase "evidence for." What I deny is that it characterizes the *only* important sense of that phrase: it is perfectly obvious that subjects who know p can gain more evidence for p, in *some* good sense. Imagine that Newton is conducting experiments to test his hypothesis that $F = MA$. Consider the moment when, by the invariantist's lights, he has just barely come to know that this hypothesis is true. Surely the centuries of subsequent physics experiments can, in some good sense, provide more evidence that $F = MA$.

What good sense is that, exactly? I think there are several. First, those who accept the Evidence-Probability Link and $E = K$ often grant that you should not be *absolutely certain* of every proposition that has epistemic probability 1 for you. It is rationally permitted, perhaps even rationally required, to harbor some doubt about some propositions with epistemic probability 1 (Williamson 2001, p. 213). Thus, here is one option:

> The *Evidence-Certainty Link*: If learning E increases your rational confidence in p, then, *in one good sense of the phrase*, E constitutes more evidence for p.

Another idea is this. Consider a pillar that holds up a floor all by itself. The floor can still gain further support from a redundant load-bearing pillar. Similarly, suppose you have some background information i which leaves open the truth of p. You then come to learn q, which entails p. Still, you might gain further evidence for p by learning r, which, *setting q aside*, raises the probability of p.

---

[17] I owe this reasoning to an anonymous referee.

Thus, even if $E = K$, there are several good senses in someone can already know p, but still accumulate further evidence for p.

Conclusion. The Epistemic Theory of Mind works best with gradable generalizations, not binary ones. However, even many gradable generalizations are not good enough. The invariantist offers us gradable generalizations that draw lines in arbitrary places. Only the contextualist gives us gradable generalizations that follow the contours of the phenomena.

Having made a thorough examination, we can at last conclude that the argument from gradability is sound. We should be robustly confident in its conclusion – again:

> AG3. If the Epistemic Theory of Mind is correct, then there is a strong reason to prefer the Contextualist Theory of Mind to the Invariantist Theory of Mind.

9. The broader significance of the argument

Suppose that the argument from gradability is sound. What broader significance does that have? In this section, I partly answer that question. First, I consider an objection, not to the soundness of the argument, but to the significance of its conclusion. After rebutting this objection, I clarify precisely which forms of contextualism are supported by the argument.

An objection. The Contextualist Theory of Mind appeals to countless *knowledge-like relations*, partially ordered by safety, certainty, and reliability. You might wonder whether we can do without the knowledge-like relations: why not just explain the subject's behavior in terms of how safe, certain, and reliable her *true beliefs* are?[18] The idea is to adopt:

The *Gradable Doxastic Theory of Mind*:

> (+) If a subject truly believes p, then: the safer, more certain, and more reliably formed her belief is, the more persistent, risk-tolerant, etc. she will be in acting on p.
>
> (−) Knowledge is not part of the best overall framework for explaining behavior at the psychological level.

I will assume that this theory draws on notions of safety, certainty, and reliability that are prior to the notion of knowledge. (Otherwise (+) would be incompatible with (−).)

The advocate of the Gradable Doxastic Theory of Mind is not objecting to the argument from gradability itself. She grants that the argument is sound. She thus grants its conclusion – that if the Epistemic Theory of Mind is correct, then there is a strong reason to prefer the Contextualist Theory of Mind to the

---

[18] I owe this objection to an anonymous referee.

Invariantist Theory of Mind. Her objection is that this conclusion is *insignificant*: the Epistemic Theory of Mind is *not* correct.

I think this objection fails. For, on the one side, the Contextualist Theory of Mind can help itself to the full explanatory power of the Gradable Doxastic Theory of Mind: assuming that generalizations like (+) are useful in the first place, the contextualist can simply take them on board.

However, on the other side, the Gradable Doxastic Theory of Mind cannot be granted much of the explanatory power of the Contextualist Theory of Mind. This is shown by the original arguments for the Epistemic Theory of Mind. For instance, as Pavese argues, any non-epistemic theory of mind cannot metaphysically explain agentive control, because this requires, not just *true belief*-how, but *knowledge*-how. Or compare two versions of Williamson's burglar. *Burglar A* stands in a fairly stringent knowledge-like relation to the proposition that the diamond is in the house, and is thus very confident that the diamond is in the house. *Burglar B* has an equally confident belief that the diamond is in the house, but this belief is based purely on a very confident *but false* belief that the diamond is in the study. The Doxastic Theory of Mind predicts that these burglars will tend to be equally persistent and risk-tolerant when ransacking the house. However, as Williamson observes, this is unlikely to be the case. After ransacking the study and learning that it is disappointingly diamond-free, Burglar B is more likely to give up; Burglar A is more likely to persevere.

The underlying problem is that the Gradable Doxastic Theory of Mind is insensitive to whether there is an *appropriate connection* between the safety, certainty, or reliability of a true belief in p, and p itself. The Contextualist Epistemic Theory of Mind is, rightly, sensitive to this.

Thus the objection is blocked: the Gradable Doxastic Theory of Mind cannot replace the Contextualist Theory of Mind.

How the argument bears on different contextualist theories. Contextualist theories can posit many different kinds of contextualist parameters on knowledge. The potential parameters that I have focused on, such as safety, certainty, and reliability, are all gradable. However, other potential parameters need not be gradable. Thus, consider two kinds of contextualist theories:

> *Partly gradable contextualism*: In different contexts, "know" refers to knowledge-like relations that differ with respect to certain parameters, where *only some* of these parameters are gradable.

> *Wholly gradable contextualism*: In different contexts, "know" refers to knowledge-like relations that differ with respect to certain parameters, where *all* of these parameters are gradable.

The argument from gradability is neutral between these theories. We are welcome to adjudicate between them by using familiar methods, such as appeals to linguistic data. However, consider this third kind of contextualist theory:

Non-gradable contextualism: In different contexts, "know" refers to knowledge-like relations that differ with respect to certain parameters, where *none* of these parameters are gradable.

The argument from gradability undermines this kind of contextualist theory, for it is incompatible with the Contextualist Theory of Mind that I have been developing.

Conclusion. The argument from gradability does, indeed, have the significance that it appears to have: the Gradable Doxastic Theory is not a suitable replacement for the Contextualist Epistemic Theory of Mind. In addition, the argument from gradability provides equal support to partly and wholly gradable versions of contextualism (although it undermines non-gradable versions).

10. Conclusion

If we wish to explain behavior in terms of knowledge, then we should be partly or wholly gradable contextualists, not invariantists: that is the central lesson of this paper. This result is unsurprising once we understand the structure of the underlying metaphysics. According to the Epistemic Theory of Mind, the properties *doing* the explaining include safety, certainty, reliability, degree of evidential support, and true belief. And the properties *getting* explained include persistence, risk-tolerance, immunity to misleading counter-evidence, proportion of true beliefs, success in navigating the environment, and agentive control. Most of the properties on the explaining end, and all of the properties on the explained end, are smoothly gradable: this is the deep metaphysical structure of things. That is why we need the explanatory ramps of contextualism rather than the explanatory cliffs of invariantism.

REFERENCES

Bricker, A. (2021). "Knowledge is a mental state (at least sometimes)." *Philosophical Studies* (online): 1-21.

Fantl, J. and McGrath, M. (2009). *Knowledge in an Uncertain World.* Oxford: Oxford University Press.

Kornblith, H. (2000). "The contextualist evasion of epistemology." *Philosophical Issues* 10: 24-32.

Kornblith, H. (2002). *Knowledge and Its Place in Nature.* Oxford: Oxford University Press.

Lewis, D. (1983). "New work for a theory of universals." *Australasian Journal of Philosophy* 61 (4): 343-377.

Nagel, J. (2010). "Epistemic anxiety and adaptive invariantism." *Philosophical Perspectives* 24: 407-435.

Nagel, J. (2013). "Knowledge as a mental state." In *Oxford Studies in Epistemology, Vol. 4*, eds. T. Gendler and J. Hawthorne (Oxford: Oxford University Press), pp. 273-308.

Pavese, C. (2021). "Knowledge and mentality." *Philosophical Perspectives* 35 (1): 359-382.

Sider, T. (2011). *Writing the Book of the World.* Oxford: Oxford University Press.

Stanley, J. (2005). *Knowledge and Practical Interests.* Oxford: Oxford University Press.

Turri, J. (2010). "Epistemic invariantism and speech act contextualism." *Philosophical Review* 119 (1): 77-95.

Williamson, T. (2000). *Knowledge and Its Limits.* Oxford: Oxford University Press.

Williamson, T. (2005). "Contextualism, subject-sensitive invariantism and knowledge of knowledge." *Philosophical Quarterly* 55 (219): 213-235.