Invariantism, contextualism, and the explanatory power of knowledge

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 will figure importantly in 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 “knows.” I think that if we accept the framework at all, we have a very compelling reason to accept contextualism instead. That is what I will argue here.

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

Behavior can be explained at many levels: at the level of physics, chemistry, neurology, 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 will figure importantly in 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 neurological or chemical. 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 (forthcoming, 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 figures importantly in some kind of overall framework for explaining behavior at the psychological level. Which framework, though? Presumably the most interesting thing to talk about is the *best* overall explanatory framework. Surely these theorists would be disappointed if knowledge figured importantly in some pretty good overall explanatory frameworks, while being entirely absent from the *very best* one.

For these reasons, I will understand the thesis that knowledge is a mental state, as affirmed by Williamson, Kornblith, Nagel, Pavese, and Bricker, as equivalent to the Epistemic Theory of Mind. But since I myself find it easy to lose track of what is meant by talk of “mental states,” I will stick to the more explicit language that I have been using.

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 “knows” refers to a single, invariant relation no matter the context. *Contextualists* disagree; they think that “knows” is context-sensitive, referring to different relations in different contexts. This gives us two very different ways of understanding the Epistemic Theory of Mind:

The *Invariantist Theory of Mind:* The word “knows” refers invariantly to a single relation, knowledge. This relation will figure importantly in the best overall framework for explaining behavior at the psychological level.

The *Contextualist Theory of Mind:* The word “knows” can, depending on the context, refer to countless different knowledge-like relations. Many or all of these relations will figure importantly in the best overall framework for explaining behavior at the psychological level.

What I find notable is that none of the theorists who accept the Epistemic Theory of Mind have committed to contextualism. By contrast, many of them, including Williamson (2005), Kornblith (2000), and Nagel (2010), have expressed at least strong sympathy for invariantism.[[1]](#footnote-1) I think that this is a mistake.

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

My thesis is that if we accept the Epistemic Theory of Mind, then we have a powerful reason to prefer contextualism to invariantism. In this section, I paint my argument for this conclusion 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 it is true, as I have heard it said, that being tall tends to make a person more attractive, but also tends to make them die younger. How can we best incorporate generalizations such as these into an explanatory framework that leverages the property of being tall? Well, notice that the higher we set the bar for tallness, the fewer instances of the generalizations we will get; but the lower we set the bar, the more exceptions our generalizations will have. This might inspire us to look for the single minimum height property that strikes the best balance between these desiderata – 73.29 inches, it turns out. Our explanatory framework would then consist of just a pair of *binary generalizations*: first, that being tall, understood as coming to at least 73.29 inches, makes people more attractive, and second, that being tall makes people die younger. This approach, which we can call the *invariantist theory of height*, has substantial explanatory power.

But we might take a very different approach. We might not bother trying to find the single minimum height property that would give us the most explanatory bang for the least explanatory buck. Instead, we might develop *gradable generalizations*: we might explain the *degree* to which someone will tend to be considered more attractive, and the *degree* to which they will tend to die younger, in terms of the *degree* of tallness that they exhibit. This approach, call it the *contextualist theory of height*, is a bit more complex than its invariantist counterpart – but also *much* more powerful. The invariantist theory of height can only explain differences between individuals whose heights are on opposing sides of the 73.29-inch mark. The contextualist theory of height not only explains that, it also explains differences among individuals who are all below the line, or all above it. It explains why being 73 inches tall makes someone more attractive than being 72 inches tall, which makes someone more attractive than being 71 inches tall, and so on; and it explains why being 74 inches tall makes someone tend to live longer than being 75 inches tall, which in turn makes someone tend to live longer than being 76 inches tall.

To my mind, the contextualist theory of height is *obviously* the better of the two. For there are many, many actual cases in which height differences correspond to differences in attractiveness and longevity, *even though* all of the individuals being compared are on the same side of the 73.29-inch line. In turn, a large part of what undergirds these patterns is that the properties to be explained – attractiveness and lifespan – are gradable ones: a subject can be more or less attractive, and more or less long-lived. The lesson is that, very often, gradable properties will be more aptly explained by gradable generalizations than by binary ones.

With this in mind, return to the Epistemic Theory of Mind. As we will see below, the advocates of this framework disagree about exactly which properties are distinctively explained by knowledge – but they all think that these properties, whatever they are, are gradable. If we combine the Epistemic Theory of Mind with invariantism, then the natural approach is to explain these gradable properties in terms of binary generalizations, e.g., “Knowing that *p* will make a subject more persistent than not knowing that *p*.” If we conjoin the Epistemic Theory of Mind with contextualism, we instead get *gradable generalizations*, e.g., “Standing in more and more stringent knowledge-like relations to *p* will make the subject more and more persistent.” These generalizations are a bit more complex than their binary counterparts, but they are enormously more powerful. That makes the gradable approach the clearly superior one overall.

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

1. If the Epistemic Theory of Mind is correct, then what knowledge distinctively explains are non-maximal degrees of gradable properties.
2. 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.
3. 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.

That is what the argument from gradability looks like from a distance. Let us inspect it more closely.

3. In defense of A1

Begin with:

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

In support of A1, 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 knowledge is, in a particular but representative example, better at explaining behavior than 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.[[2]](#footnote-2)

 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 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 counter-evidence, a moderate amount, or a lot. The burglar’s knowledge surely does not immunize him to any amount of counter-evidence whatsoever. It immunizes him to a particular, non-maximal amount of counter-evidence.

 *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.[[3]](#footnote-3)

 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 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 would like to pause to register a concern: I worry that Kornblith’s theory entails that knowledge *exists*, but it does not entail that knowledge actually *explains* 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 reliably produced true beliefs, i.e., many instances of knowledge. But what explanatory work is knowledge doing, exactly? 3 does not explain 2 by appealing to *beliefs* that are reliably produced and true, but by appealing to *cognitive capacities* that reliably produce true beliefs – the reliability attaches to the capacities, not the beliefs. And 2 does not mention reliably produced true beliefs at all. Thus, even taken on its own terms, Kornblith’s theory seems to leave knowledge with no explanatory role. For this reason, I happen to prefer Williamson’s argument for the Epistemic Theory of Mind.

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? They certainly are. 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. For all their knowledge, 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*: agentive control 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.[[4]](#footnote-4)

Does agentive control come in degrees? Of course it does. An expert pool player will have a very high degree of 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 A1: 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. In defense of A2

The remaining premise to consider is:

1. 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 particular explanatory approaches that we wish to compare.

 *Two versions of the Epistemic Theory of Mind*. For the invariantist, the obvious way to develop the Epistemic Theory of Mind is to develop binary generalizations that appeal only to a single, invariant relation of knowledge. There is room to bicker over the precise form of these generalizations. One option is: knowing *p* makes the subject instantiate property *F* (e.g., persistence) to at least degree *d* when she acts on *p*. Another option is: knowing *p*, as compared to not knowing *p*, makes the subject have property *F* to a significantly higher degree when she acts on *p*. I am sure you can think of other options, too. The details will not be important for now.

What *is* important is that this invariantist approach is not the only way of developing the Epistemic Theory of Mind. There is a natural contextualist alternative, which is to appeal to countless *knowledge-like relations*. Crucially, although the invariantist will deny the explanatory relevance of all but one of these relations, even she should concede that these relations all *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 these 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*-.

You might feel the tickle of a worry here. The contextualist begins with various gradable properties (safety, etc.) that are supposed to accompany knowledge, and tries to partially order them in terms of stringency. Does this maneuver presuppose that knowledge-like relations can be *analyzed* in terms of true belief plus these gradable properties? If so, then we should resist the idea that there is any such stringency ordering: analyses of knowledge have not fared well, and we have no reason to be more optimistic about the prospects of analyzing knowledge-like relations more generally.

Fortunately, the worry can be dispelled: we can make our partial ordering without assuming that knowledge can be analyzed. By analogy, take all of the shades of red, which can be partially ordered in terms of saturation and brightness. This does not presuppose 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.

With our partial ordering in hand, we can build a new explanatory approach. Where the Invariantist Theory of Mind posits binary generalizations about knowledge, the Contextualist Theory of Mind gives us gradable generalizations about knowledge-like relations. Here is a representative one: standing in more and more stringent knowledge-like relations to *p* makes the subject more and more persistent (or risk-tolerant, or …) when she acts on *p*.

*The theories compared*. At last, we are in a position to evaluate A2, the claim that 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. Suppose 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 smoothly to other gradable targets.) How would the Contextualist and Invariantist Theory of Minds compare to each other?

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 she 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 difference 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 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 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.

As an incidental bonus, the contextualist can even explain the 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, it is worth investing in gradable generalizations about the explanatory power of knowledge.

There is a distinction between two forms of invariantism that might seem to complicate things. 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.[[5]](#footnote-5) 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?

I think the answer is no. All that the intellectualist and the 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 smooth gradients in behavior, and *no* single line can possibly do that, no matter where you put it. To express things another way, the subject-sensitive invariantist agreeswith her intellectualist counterpart in positing only binary generalizations; it is just that 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.

On these grounds, I find A2 very compelling. That, in turn, completes our initial defense of the argument from gradability: we can tentatively conclude that there is a strong reason to prefer the Contextualist Theory of Mind to its invariantist counterpart. But before we render our final verdict, let us inspect the argument more carefully.

5. Objections and replies

Here is the argument again:

1. If the Epistemic Theory of Mind is correct, then what knowledge distinctively explains are non-maximal degrees of gradable properties.
2. 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.
3. 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, premise A1 is watertight; it just expresses the commitments of those who defend the Epistemic Theory of Mind. Premise A2 is where the action is. Here the natural move for the invariantist is to try to replace her original binary generalizations with more nuanced ones while still appealing to just a single relation of knowledge. Several options present themselves.

*First option*. It is common for invariantists to appeal not just to knowledge, but also iterations of it.[[6]](#footnote-6) 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* makes 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 humans. 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 be 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 (ii) someone who truly believes *p* on the basis of substantial evidence while falling barely short of knowledge 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 differ only slightly. These are exactly the same kinds of problems that arise for the original version of the Invariantist Theory of Mind. The root cause of the problem is that these approaches give us a stair-like explanatory structure when we need a smooth ramp.

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

*Second option*. The invariantist might suggest this pair of gradable generalizations. First, that the closer a true belief is to knowledge, the more persistent, risk-tolerant, and so on the subject will be. Second, that the further “past the post” of knowledge a true belief gets, the more persistent, risk-tolerant, and so on the subject will be. Once this is done, the invariantist might hope, her framework will have 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 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 that goes much deeper: the line between knowledgeable and non-knowledgeable true beliefs has become completely 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.

*Third option*. It is sometimes held that a subject’s evidence is identical to what she knows: E = K.[[7]](#footnote-7) 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 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. To be fair, however, the invariantist still 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, 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 be.

The invariantist’s idea was to explain how persistent (e.g.) Mata Hari is 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 the lights of the current proposal 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 is 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 explain at all 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 large and smoothly gradable variations in the subject’s persistence, risk-tolerance, and immunity to misleading counter-evidence. The invariantist approach draws only binary distinctions, ones that are inadequate to the explanatory task. The contextualist approach gives us the gradable distinctions that we need.

*Fourth option*. The invariantist might try expressing her generalizations in terms of probabilistic knowledge. Such knowledge can take forms like these: *probably p*, *perhaps q*, or *r is more likely than s*. The invariantist might then commit to generalizations like this: the more likely a subject knows *p* to be, the more persistent she will be when acting on *p*.

*Reply*. By themselves, such generalizations are not enough: they do not explain the 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 be radically different in, for instance, their immunity to misleading counterevidence. The current approach leaves us powerless to explain these large differences.

*Fifth option*. The invariantist’s third option was to say that the more evidence – understood as invariant knowledge – the subject has in support of *p*, the more persistent she will be. The problem was to explain differences among subjects who have no further evidence for *p*. The fourth option was to say that the more likely the subject knows *p* to be, the more persistent she will 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 they know *p* to be. And she can explain the differences among subjects who all know the same proposition *p* by observing that these subjects can still have much more evidence for *p* than others.

*Reply*. As far as I can tell, this is the best way for the invariantist to develop the Epistemic Theory of Mind. Still, after careful reflection, I think that the contextualist approach proves superior. The basic reason, to my mind, 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: once in a while pretty 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.

To be sure, the invariantist can patch things up: she can add an exception for cases in which one generalization applies *because* the other generalization applies. 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. What we want is a gradable generalization that explains these differences.

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.

*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 worked through these ideas, we can at last conclude that the argument from gradability is sound. We should be robustly confident in its conclusion: if we accept the Epistemic Theory of Mind, then there is a strong reason to prefer contextualism to invariantism.

6. Conclusion

If we wish to explain behavior in terms of knowledge, then we should be contextualists, not invariantists: that is the central lesson of this paper. This result is unsurprising once we recognize 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.

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. (forthcoming). “Knowledge and mentality.” *Philosophical Perspectives*.

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.

1. 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.” [↑](#footnote-ref-1)
2. For the complete argument, see Williamson (2000, pp. 60-64). [↑](#footnote-ref-2)
3. See Nagel (2013, especially §3-§4); Bricker (2021, especially §4). [↑](#footnote-ref-3)
4. See Pavese (forthcoming, especially §3 and §7). [↑](#footnote-ref-4)
5. Intellectualist invariantism is the traditional view. For some detailed defenses of subject-sensitive invariantism, see Stanley (2005); Fantl and McGrath (2009). [↑](#footnote-ref-5)
6. See for example Williamson (2005, p. 231); Turri (2010, p. 87). [↑](#footnote-ref-6)
7. See Williamson (2000, ch. 9). [↑](#footnote-ref-7)