Understanding Attribution: Problems, Options, and a Proposal

Felipe Morales Carbonell

Pre-print version

Please cite the published version at

https://onlinelibrary.wiley.com/doi/epdf/10.1111/theo.12380

Abstract

In this paper, I give an overview of different models of understanding attribution, and advance a contextualist account of understanding attribution. While other contextualist accounts make the degree in which the epistemic states of the relevant agents satisfy certain invariant conditions context-sensitive, the proposed account makes the conditions themselves context-sensitive.

1 Introduction

The minimum that a general account of understanding must do is to provide an answer to the following two questions:

1. When is a subject attributable with understanding?
2. How do subjects compare with regards to their understanding?

Call the first question the *outright attribution* problem. In a semantic key, it is equivalent to the problem of providing the truth conditions for statements of the form

S understands [X]

where [X] stands for what is understood (the object or target of understanding). Examples are easy to come by:

3. Susan understands the operation of the machine
4. Andreas understands the feelings of her mother,
5. Stephan understands why his relationship with Anne failed

and so on.

Call the second question the *comparative judgment* problem. In a semantic key, it is equivalent to the problem of providing the truth conditions for statements of the form

S_1 understands [X] [better/more/…] than S_2

where the type of comparison might be variable. Some examples:

6. Susan understands the operation of the machine better than Angela,

7. Susan understands better than before,

8. Susan understands more than before

These problems are closely related. [1] A common view about understanding states is that they are a) valuable, and b) valuable to various degrees. Understanding attribution seems to track the evaluative properties of the states of the relevant agents. Already Rosenberg (1981), in an early paper on understanding, argued that attributions of understanding have as their main function the signaling of their subjects' positive epistemic status. More recently, Hannon (2019) has defended the view that we attribute understanding why to good explainers, and Hills (2009, 2016) has linked understanding attributions with certain kinds of expertise or skill. In all these cases, understanding attributions establish relationships of epistemic deference towards the individuals that are deemed attributable with understanding. These views seem to be in accord with seemingly deeply entrenched practices in ordinary language (cf. Wilkenfeld, Plunkett, and Lombrozo (2016)).

These observations had led many authors to draw the conclusion that this connection to positively evaluable states is an important part of the *nature* of understanding. In effect, many authors (like Pritchard (2009), to give a well known example) argue that understanding just *is* a type of cognitive achievement. The debate on the metaphysics of understanding has tended to focus, consequently, on the question whether understanding satisfies one or the other success conditions (like factivity, safety, grasp, clarity, and so on). These properties have been taken as constitutive of understanding states *qua* successes.

There are some subtleties that are easy to miss here. It is theoretically possible for us to have a theory of understanding that abstains from saying that we should conceive of understanding (and the states to which it corresponds) as a form of achievement, and yet have attributions of understanding track the conditions that mark certain states as successful in some relevant sense. In those cases, the constitutive conditions of being an understanding state and the constitutive conditions of being a successful understanding state come apart–plausibly, successful understanding states could be a proper subset of all understanding states. It might be possible, for example, that subjects who make understanding attributions only do so for a specific subclass of understanding states–for some conception of the latter. [2]

---

[1] Compare with the setup of Khalifa (2017) and Baumberger (2019). Both authors center their approaches to attribution on the question of how understanding can be gradual. Khalifa also introduces the issue of generic attributions of understanding (p. 5), but this is secondary.

[2] Indeed, I think that there are reasons to think that something of the sort is the case, although in this paper I won't lay out the full view. The important thing is that the models of attribution should be able to make room for this kind of view, and this is something I try to pay attention to in the proposed models.

Thus, it might turn out that the way in which understanding attributions are made by ordinary speakers only partially matches the underlying structure of the relevant states of the agents attributed with understanding. Ordinary understanding attributions indicate only indirectly a path to the correct metaphysics of understanding, and we cannot read ontological commitments straight off understanding attributions. Recent empirical work on attributions (like a recent series of papers by Daniel Winkelfeld, Tania Lombrozo and Dillon Plunkett (2016, 2018)), even though of intrinsic interest, will have a limited payoff when it comes to these metaphysical questions–unless we ought to accept some form of metaphysical picture that minimizes the gap between ordinary language and metaphysics, perhaps in a deflationist spirit. While I favor a form of deflationism myself, I think we ought to deflate the metaphysical picture through the application of a methodology of explication or conceptual engineering, that can preserve the gap between the ordinary language and metaphysics.[3] Once we construct a rich enough picture or model of our target (in this case, of understanding), we should be able to explain the utterances of ordinary speakers. For that purpose, it will be necessary to have empirical studies of the patterns of attribution that the theory needs to explicate. In any case, here I will try to limit myself to the part of that picture that would describe the mechanisms of attribution, and nothing else.[4]

Here, I have three goals. First, I want to put together a framework to think about understanding attribution–a way to see how the different theoretical solutions to the problem of understanding attribution can be arranged in logical space. I think this is necessary because existing accounts usually are not detailed enough to have a fully concrete picture of what they are supposed to entail. The classification key that I will use is the ways in which theories of understanding can account or fail to account for the sensitivity of understanding attributions to contextual factors. Second, I intend to sketch a novel account of understanding attributions. I will argue that we should take understanding attributions to be thoroughly context-sensitive. In more common context-sensitive views, only something like a threshold shifts with the context. In the proposed model, it is the type of measures that are relevant for attribution that is determined by the context. Indeed, we might even want to introduce elements of a relativistic story in order to accommodate cases of faultless divergence between understanding agents in a single conversational context, which other theories fail to acknowledge or handle incorrectly. Third, and finally, I want to examine the methodological consequences about the metaphysics of understanding that follow from the results herein obtained. I want it to be clear at the outset that, more than wanting to convince the reader of my own proposed solution to a series of problems, I want to raise awareness about the richness of potential difficulties and solutions that the general questions raised bring with them when thought about at length.

I will proceed as follows. In section 2 I will explore the logical space of solutions to the outright attribution problem. In section 3 I will elaborate on what I think should be the preferred model of understanding attribution, along the lines I just indicated. Then, in section 4 I will compare the proposed model with the models (both explicit and implicit) in various theories of understanding from the literature.

---

[3] Baumberger (2019) also uses a methodology of explication, although he does it in a more traditional way. I will discuss his account in section 4 below.

[4] The kind of modeling I will engage in is what Weisberg (2013) calls *generalized modeling* (cf. ch. 7). It is also similar to the construction of what Strevens (2004) calls a *framework*.

2 An overview

In this section, I will give an survey of the solution space to the outright attribution problem. The key to understand the problem is the way in which different accounts handle the apparent context-sensitivity of understanding attributions. Like in the literature on knowledge ascriptions, the solution space ranges from bare, context-insensitive, invariantist positions (2.1), to context-sensitive invariantist views (2.2), to contextualist solutions (2.3), to relativist approaches (2.4). I should warn at the outset against some possible misunderstandings with my use of this terminology. I am grouping these views according to whether a function of evaluation of states shifts with the context. Accordingly, I will classify as properly contextualist only those views where that function (or whatever plays that role) can shift.[5] If we rather characterize contextualist as the view that the appropriateness of outright attributions depends in important ways on the context, what I call context-sensitive invariantist views should count as contextualist already–of a weaker sort.[6] I will show how these different solutions build on the elements of simpler solutions within this solution space, forming a hierarchy of classes of increasing complexity.

It should be noted that of many many concrete proposals to the outright attribution question in the literature tend to cluster together in the context-sensitive invariantist camp (for the reasons I have just given, proponents of those accounts would not necessarily agree with this label). This trend leaves some important gaps open–to my knowledge, for example, there has been no previous attempts at producing concrete proposals along contextualist or relativist semantics for understanding attributions like those I sketch here.[7] I will attempt to fill in the resulting gaps by sketching some of the missing potential solutions. On the other hand, some of the invariantist positions I will sketch here probably have no explicit defenders in the literature. The reason to include them is that they illustrate how a theory of understanding attribution can be built in stages, and this allows me to introduce concerns at a more relaxed pace for those readers who are not necessarily up to speed with the relevant philosophy of language.

As a final caveat, it must be kept in mind that I will remain neutral on whether understanding reduces to knowledge.[8] This will make things slightly harder for myself, since otherwise I could have simply lifted the models from the relevant proposals from the literature. However, I do this in part because I don't think that such reduction would succeed (which I won't defend here), but most importantly because I think that all theories of understanding should provide some solution to the problem of understanding attributions, and the resources available to do this do not discriminate between intellectualist and anti-intellectualist positions. In any case, I will defer a defense of my own preferred solutions to the problem for section 3. In this section I will limit myself to map out the land.

2.1 Context-insensitive invariantism

I will make the assumption that understanding attributions are linked to the epistemic value of the state of the subjects of attribution. More specifically, I will assume that the appropriateness of attributions is a function of the given values of those states. I will further assume that the

[5] For various ways in which we can carve the solution space here, cf. MacFarlane (2011).
[6] MacFarlane would call them 'nonindexical contextualists'. Cf. MacFarlane (2009).
[7] There are some antecedents, like Carter (2014), who considers the possibility of a relativist semantics for understanding-why, but those are underdeveloped.
[8] Or really, any other substantive question about the nature of understanding.

value of a subject's epistemic state is grounded on the state's satisfaction of some conditions (for example, one's epistemic state might have value only insofar as it is intelligible, or truthful, and so on). It is thus possible to formulate the requirements for the appropriateness of attribution either in purely evaluative terms, purely in terms of the satisfaction of the underlying conditions or in terms of a combination of both. What we are interested in here is the *form* of the conditions for attribution, so in what follows I will use a schematic formulation of those requirements except to give examples.

The simplest models of understanding attribution single out the satisfaction of one single condition across all contexts as a requirement for attribution–a form of invariantism. In their most bare form:

Bare Invariantism

There is a single function P, which maps a subject s' states to truth values across all contexts, such that a subject s is appropriately attributed understanding at a time t iff s's state e at t satisfies P.

A state e satisfies P iff Pe is true. We can extend this to capture attributions of understanding with specific objects or targets if we treat P rather as a relation between the subject and some object: a subject s is appropriately attributed understanding of o at a time t iff s's state e at t satisfies sPo (where sPo is satisfied if it is true that s stands in the relation P to o). In the overview I will omit this complication, but I will return to it when I give my own proposal in section 3.

For example, someone could have the theory that understanding is a form of grasp. If they defended Bare Invariantism, their account of attribution would look like

Grasp Bare Invariantism

S is appropriately attributed with understanding X iff S is in a grasping relation to X[9]

To give a different example, suppose that we had a theory of understanding in terms of intelligibility. Then, this could be a Bare Invariantist version of their conditions for outright attribution:

Intelligibility Bare Invariantism

S is appropriately attributed with understanding X iff X is intelligible for S[10]

These models needs to be amended in order to make sense of the idea that understanding comes in degrees. This is important if one wants to have a somewhat unified account of outright attributions and comparative judgments (with the caveats given above). One way to do this is to adjust P to be a function that gives values that can be identified with degrees (that is, we change the range of the function from $\{\text{true}, \text{false}\}$ to a range $\{v_0 \ldots v_n\}$, which is intended to be well

[9] Cf. Strevens (2020). Khalifa (2017) also explicates understanding in terms of grasp.

[10] Cf. De Regt (2017), although there intelligibility is not the only dimension of understanding.

ordered).[11] I will call these 'epistemic values'. A state e satisfies P adjusted in this way iff its value is non-zero/not the least-value. We adjust Bare Invariantism accordingly:

Bare Invariantism*

There is a single function P, which maps a subject s' states to epistemic values across all contexts, such that a subject s is appropriately attributed understanding at a time t iff s's state e at t satisfies P.

With this amendment, it is possible to explain the gradability of understanding in terms of range of values of the P function.[12] We might want to say that the degree of an agent's understanding is the epistemic value of their satisfaction of the P function. This allows giving a unified model of outright attribution and some comparative judgments. A natural proposal is the adoption of

Simple Comparative Judgment

Given the function P that assigns epistemic values to states, S_1 understands [better/more] than S_2 iff $P(s_1) > P(s_2)$

Take the case where two agents, satisfy the P function to varying degrees. Then we can say that one of them understands better, or more, than the other if their degree of understanding is greater.

Despite the elegant simplicity of the suggestion, it cannot work. Two problems arise immediately. First, because satisfaction of P is sufficient for the truth of outright attributions, Bare Invariantism coupled with Simple Comparative Judgment cannot make sense of comparisons between subjects where at least one of them does not satisfy the conditions for outright attribution. Second, the model is not rich enough to make sense of comparisons along different evaluative axes–in fact, it does not have enough resources to even handle the difference between 'understanding more' and 'understanding better'.

A solution to the second problem at least requires both modeling epistemic values differently and giving a different account of comparative judgments. This can be done in several ways; for simplicity I will only sketch one way to do it. First, we recognize that the epistemic state of an agent can be evaluated along a multitude of dimensions.[13] The epistemic value of the state can be seen as the *aggregate* of these different values. One way to represent this is as an n-tuple, where n is the number of dimensions of evaluation that we want to consider (the set of dimensions of evaluation will come from a 'concrete' theory about understanding).

Take again the example of the theory of understanding as grasp. Suppose that grasp can be evaluated in terms of its scope (how much is grasped) and intensity (how firmly its object is

[11] I will assume in what follows that we can assign real numbers to the degrees, but this is only for simplicity of exposition. There are more sophisticated models in the literature on the semantics of comparatives. Cf. Pinkal (2005), who uses atomic degrees, Schwarzschild and Wilkinson (2002), who use intervals, and Moltmann (2009), who uses tropes.

[12] Khalifa (2013) and Kelp (2015) have suggested, in line with this, that degrees of understanding should be thought of in terms of the distance of states from an ideal of understanding in terms of the characteristics of such an state.

[13] In Kvanvig's (2009) 'quasi-factive' view, understanding is evaluated in terms of their truthfulness and 'centrality'. In Kelp's (2015) account, understanding has to be both knowledgeable and well-connected to satisfy the relevant grading function (USKWC). Wilkenfeld (2015) takes accuracy and intelligibility as the significant scales for evaluating understanding (USAI), and suggests that in some cases there might be even more dimensions to consider in the aggregate. Cf. also Bengson (2018) and Baumberger (2019).

grasped). Call this the *2D grasp theory* or 2DG for short. [14] Suppose that there is a couple of functions, S and I, that take the state of an agent and give numeric values to the scope of their grasp (about some object) and the intensity of that grasp (the values are normalized so they fall under the same scale–this simplifies comparisons later). Now, we make the function P build, for any state e, the tuple $\langle S(e), I(e) \rangle$. With this information, we can sketch the following semantics for comparative judgments. Whether an agent understands more than another seems to be a matter exclusively about the scope of their grasp. Whether an agent understands better seems to be a function of both the scope and the intensity of their grasp. One could give the following semantic clauses for these comparisons:

Understands More for 2DG

Given the function P that assigns epistemic values to states, S_1 understands more than S_2 iff $P_S(s_1) > P_S(s_2)$ (that is, iff the scope of s_1 is greater than the scope of s_2)

Understands Better for 2DG

Given the function P that assigns epistemic values to states, and a ratio r (a real value between 0 and 1) that represents the importance of scope for the overall goodness of the state, S_1 understands better than S_2 iff

$$\left(P_S(s_1) * r\right) + \left(P_I(s_1) * (1-r)\right) > \left(P_S(s_2) * r\right) + \left(P_I(s_2) * (1-r)\right)$$

(that is, iff the weighted sum of the scope and intensity of s_1 is greater than the weighted sum of the scope and intensity of s_2)

In general, for weighted sum models of n-dimensions, we will need to specify the ratios for every n-dimension. More generally still, we could have a set of weights instead, so that the sums or aggregates are biased in different ways:

Understands Better for 2DG (Weighted)

Given the function P that assigns epistemic values to states, and set of weights $w_1 \ldots w_n$, S_1 understands better than S_2 iff

$$\left(P_S(s_1) * w_1\right) + \left(P_I(s_1) * w_2\right) > \left(P_S(s_2) * w_1\right) + \left(P_I(s_2) * w_2\right)$$

Different comparisons should be possible depending on whether we think that some specific information about the epistemic state of the agents encoded in their epistemic values can be meaningful. For example, in the context of the grasp theory of understanding it might be natural to think that we can also compare the intensity of the understanding of two agents through statements of the form 'S_1 understands more intensely/firmly than S_2'. As a general lesson, it should be clear that talking about *the* degrees of understanding is misleading: a single state of understanding can have multiple degrees of different types. While we typically speak only of the quantity and quality of understanding, this could be explained away as a contingent feature and limitation of ordinary language.

The possibility of comparisons along multiple dimensions gives some hints towards a solution for the problem of comparing states which are outright attributable with understanding to states

[14] Cf. Strevens (2020).

which are not. Suppose that the appropriateness of understanding attributions requires the satisfaction of various independent factors, each gradable along a number of dimensions. Then, it will be possible to compare states that are attributable with understanding with states that are not, but which satisfy some of the conditions for outright attribution.[15]

To illustrate this, let's suppose that our substantive theory tells us that understanding requires both intelligibility and representation manipulability, that these are independent factors, and that neither is sufficient on its own for understanding. Suppose that intelligibility is gradable along dimensions A, B, and C, and that representation manipulability is gradable along dimensions D and E. Take an agent S_1 whose epistemic value is $\langle A=.3, B=.7, C=.8, D=.1, E=.3 \rangle$ and an agent S_2 whose epistemic value is $\langle D=.9, E=.6 \rangle$. Suppose that this is sufficient for S_1 to be attributable with understanding, while S_2 is not. Now, it is possible to compare S_1 with S_2 along some dimensions (namely, D and E), and along functions of those dimensions. Of course, this does not solve in fact the problem unless those comparisons can support judgments of the form 'S_1 understands better than S_2' or 'S_1 understands more than S_2'. However, it cannot be said that in principle the model lacks the resources to provide a solution, at least for a limited range of cases.

There is one final modification one could introduce to the model. The notion of satisfaction of the conditions that support outright attributions that we have adopted so far has it that *any* positive value (or aggregate of positive values) is sufficient for understanding. But we might want to say that outright attribution is correct only of some of those values are higher than some specific *threshold*. Perhaps, just to give an example, too weak a grasp or too little ability to manipulate the relevant representations, are not sufficient for someone to be attributable with understanding. When only one measure of the agents' epistemic state is singled out as their degree of understanding, we get:

Threshold Invariantism

> There is a single function P, which maps states to epistemic values across all contexts, and a threshold value t which corresponds to a possible value of P. For a subject s to be appropriately attributed with understanding, their state e must satisfy P at least to degree t.

In the general case, when epistemic values are represented by an n-tuple of values, as described above, we will need a *threshold specification*: an n-tuple that fully specifies the thresholds for all

[15] The model can only support comparisons if it is applied to a theory of understanding whith more than one contributing independent factor. For example, it cannot do this for the grasp theory of understanding. Since grasp itself is sufficient for understanding, and the measures are dependent on there being a grasp relation, we cannot compare something that has a value for some of the measures against something that does not have it: if the relata for comparison have one of the measures, they must have values for all of the measures. Rather than assigning a value of 0 to the relevant measures to incomparable states, we should disallow the comparisons as a category mistake. In practice, however, there will be some confusion on whether the measures are in effect comparable. For suppose that one tried to compare a state that can be measured in terms of quantity with something that also can be measured in terms of quantity. The relevant notion of quantity itself might vary depending on the kind of things that it ranges over. Unless this matches, the comparisons will not be legitimate. Someone could ask, for example, how somebody who has many personality defects compares against someone who has many properties–but that would be nonsense.

the relevant values in a prototypical state.[16] We adjust the condition for outright attribution as follows:

Threshold Invariantism*

There is a single function P, which maps a subject s' states to epistemic values across all contexts, and a threshold specification T, such that a subject s is appropriately attributed understanding at a time t iff for s's state e, for every value i in $P(e)$, $i > T_i$.[17]

In models like these, the explicit comparisons between $P(e)$ and T could have additional requirements, such as having different components of $P(e)$ and T weight differently. In what follows I will mostly omit discussion of these possible complications. However, I will come back to the point later, when I argue for my own proposal–one of the arguments there will be that it is unrealistic to require a single specification to apply to every context.

From a more general point of view, we have to note that these variations of the invariantist model make the conditions of outright attribution supervene on implicit comparisons between the subject that is being attributed with understanding and a hypothetical subject that meets the relevant threshold or specification minimally: a subject is attributable with understanding iff they meet the threshold equally or better than that hypothetical subject. This results in the unification of the accounts of outright attribution and of comparative judgments (beyond the fact that they supervene on the satisfaction of the same underlying conditions, which was already the case in Bare Invariantism*).

2.2 Context-sensitive invariantism

An important issue with the invariantist proposals we have considered so far is that they do not allow for any kind of context-sensitivity for understanding attributions. And yet, this is something that we would like to have, even within an Invariantist framework.[18] Context-sensitivity is simply one of the features of phenomena that we need to save–any account that does not account for it is fundamentally inadequate.[19]

If we start from a threshold model, there is a simple way to account for at least one form of context-sensitivity. Rather than assuming that one threshold is applied in every context, one can make the context fix the threshold. So while in every context the same conditions need to be satisfied, how they are satisfied can vary. For Threshold Invariantism, the adjustment is small:

Context Sensitive Threshold Invariantism

There is a single function P, which maps states to epistemic values across all contexts. For each context c there is a threshold t_c, such that for a subject s to be appropriately attributed with understanding, their state e must satisfy at least P to degree t_c.[20]

---

[16] Cf. Threshold specifications in this sense are similar in structure to what Camp (2014) calls *characterizations* (characterizations can contain more than thresholds, however).

[17] One could have more complicated versions of this where the quantification over values is different, or where there is some weighting of values (over whatever is done with the P function itself).

[18] Later I will sketch some ways in which an invariantist could do this.

[19] For a fuller discussion of why we would like to have a theory of understanding that includes context-sensitivity, see De Regt and Dieks (2005), De Regt (2009), De Regt (2017), Khalifa (2013), Wilkenfeld (2013), Kelp (2015), Bengson (2018) Hills (2016), and Hills (2018), among others.

[20] The adjustment for Threshold Invariantism* is very similar, so I will omit it here.

Alternatively, we can say that there is a function T that assigns a threshold value for every context ($T(c)=t_c$). Figure 1 gives a diagram of the model (p is the relevant attribution sentence, and A is the function that maps the triple $\langle p, T(c), P(s) \rangle$ to truth values for that sentence, with corresponds to the condition for outright attribution).

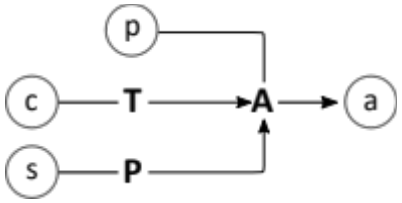

Figure 1: CSTI

A minor variation of the model specifies a *default* threshold, and *specific* thresholds only for some contexts. This can be captured by the given formulation without loss of generality (indeed, the context-insensitive version of Threshold Invariantism is also a special case of it–namely, where all contexts specify the same threshold value).

As I pointed out above, the invariantist threshold models make outright attribution supervene on implicit comparative judgments. The context-sensitivity of outright attribution could depend on the context-sensitivity of comparative judgments; this gives a second way in which understanding attributions could be context-sensitive. Take the 2DG account of understanding with the comparison model based on weights. In that account, comparative judgments depend on the weighted values of scope and intensity. We can make the weights $w_1 \ldots w_n$ be defined by the context: in some contexts intensity might matter more, and in others scope might matter more. If we allow this contextual variance for comparative judgments, even if we set a fixed threshold, in different contexts it could happen that the same state of the subject (with the same epistemic values) will be comparatively better or worse than the threshold.

2.3 Higher orders of context-sensitivity

A different way to account for the context-sensitivity of understanding attributions is to drop the assumption that a single function that determines epistemic values is at play in every context, replacing this assumption with the idea that a multitude of functions with different parameters are selected in different contexts (I will call this kind of account *Parameter Contextualism*). To illustrate the idea, we can build on Context-Sensitive Threshold Invariantism, in which case we get the following formulation:

Threshold Contextualism

For each context c there is a function P_c selected by the context which maps states to epistemic values and a threshold t_c, such that for a subject *s* to be appropriately attributed with understanding at a context k, their state *e* at k must satisfy at least P_k to degree t_k.[21]

Alternatively, we can say that there is a higher-order function P that maps contexts c to functions P_c that map states to epistemic values, such that for every context, to be appropriately attributed

[21] If we drop the condition that the functions P_c are selected with the context, we get a general schema of which the invariantist models are also a model of (they are the special cases where every context shares the same function occupying that role.)

with understanding is a way of satisfying the respective P_c function. Figure 2 gives a diagram of the model.

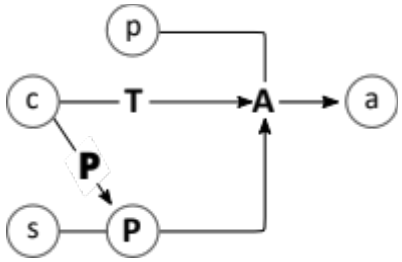

Figure 2: TC

On top of this, different substantive theories of understanding will add some constraints on what kind of P_c functions are admissible. An important case to consider is that accounts of understanding that assume some form of epistemic value monism will have to endorse the restriction that

Epistemic Value Monism Constraint

For every context c, satisfaction of a P_c will require satisfaction of a condition M that is linked to the possession of final epistemic value.

The idea is that even if as a whole the function P_c can vary from context to context, it is essential to all functions that occupy that role that they satisfy some common condition that is connected to final epistemic value. Plausibly, pluralists would want to reject the uniqueness of a function that could play that role.[22]

2.4 Relativist semantics

A different class of models for understanding attributions would make use of the framework of relativist semantics. Here I will focus only on relativist frameworks in the style of MacFarlane (2014). Above, I have been talking about 'the context' without specifying *which* context is relevant to the evaluation of outright attributions. For invariantist positions the only context that matters is the context of the subjects of attribution. Traditional contextualists would say that the context of attribution (the point at which the attributers assert the attributions) matters. According to the relativist proposal, the relevant context is the context of *assessment* (the point at which the assertions of attributers are evaluated.) This can match or not the context of attribution.

Once again, we can see the relativist model as an extension of the models we have considered so far. While before we only considered the context of attribution (in MacFarlane's terminology, the context of use), we now also need to consider the context of assessment. We are forced to give the semantics of understanding attributions in two stages. First, we obtain a representation of the state of the subject of attribution (this serves as the *content* of the attribution). Then, we assign a truth value to the attribution at the point of assessment according to some standard of evaluation. In the invariantist and contextualist models, the P functions assigned values to states

[22] It is not straightforward whether pluralist positions would want to specify any substantive restriction on this point, even though this is an open option for them. However, there might be types of constraints that any view would have to respect.

which were immediately consumed by the appropriateness function. Now, we will have a function M that gives a measure of the states, and a function P that gives an assessment of the measure of the states. Values such as thresholds and weights are given by the context of assessment.

We can illustrate these ideas with a relativist version of a threshold model:

Threshold Relativism

There is a single function M, which maps states to epistemic values across all contexts, and a function T that maps contexts to thresholds. For each context of assessment n, there is a function P_n that maps epistemic values to epistemic values. For a subject s with state e to be appropriately attributed with understanding at a context of assessment n, the value of $M(e)$ must satisfy P_n at least to degree $T(n)$.

Figure 3 gives a diagram of the model:

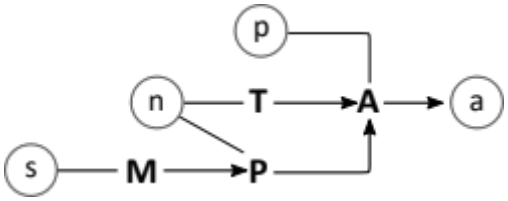

Figure 3: TR

In the contextualist proposal, outright attribution was dependent on the threshold determined by the context of use. In the current version of relativism, in contrast, nothing really depends on the context of use–in this sense, we can understand the model as an extension of an invariantist model (this is similar to how MacFarlane conceives of relativism for knowledge attributions, as a synthesis of some features of both invariantism and contextualism).

3 A proposal

Now that I have given an overview of the different types of solutions available for the problem of outright attribution, I want to present and motivate the adoption of a more elaborate account, which I will call the *Task-Based Model of Understanding Attributions* (TBMU). This model combines features of relativist and parameter-contextualist models. In section 3.1 I will sketch the model. In section 3.2 I will motivate the parameter-contextualist side of the model, and in section 3.3 I will defend its relativist side. In section 3.4 I will address the need to incorporate subject matters into the model.

3.1 The Task-Based Model

Like relativist models, the TBMU has two components. The first gives a description of the state of the subjects of attribution in some situation. The second gives an evaluation of the state thus described at a point of assessment (which is potentially different from the situation where the description is given). Unlike in the Threshold Relativism model, in the task-based model there is no assumption that the first stage will be fulfilled by a single function in all contexts. Rather, it varies with the context of the subject of attribution, like in the Threshold Contextualism model

described above. This imposes a constraint on the evaluations at the point of assessment, because assessors need to make a judgment about the relevant context of the subject of attribution besides their own.[23] This models the idea, that I will elaborate on in a second, that the contexts of assessment and the subject are in a certain relation of continuity: common to both points there is a unity of *task*. Before elaborating on this, it will be useful to formulate the model as I have done with the other models:

Task-Based Relativism

There is a function T that maps contexts to thresholds.[24] For each context c there is a function M_c selected by the context which maps states to epistemic values. For each pair of context c and context of assessment n, there is a function P_n^c that maps epistemic values to epistemic values. For a subject s with state e to be appropriately attributed with understanding at a context c and context of assessment n, the value of $M_c(e)$ must satisfy P_n^c at least to degree $T(n)$.

We can also say that there is a function **P** that assigns a function P_n^c for every pair of contexts c and n ($\mathrm{P}(c,n)=P_n^c$). This function represents the way in which the context of the task (double-indexed to the contexts c and n) selects the evaluation function.[25] A diagram of the model is given in Figure 4. In the rest of this section, I will elaborate on, and defend, the features of the model.

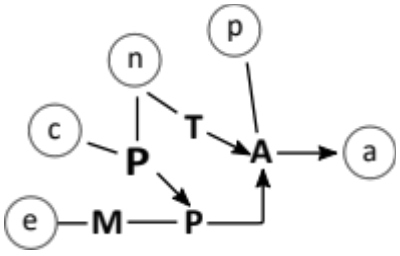

Figure 4: TBMU

3.2 Parameter-contextualism

The TBMU model just presented has two features: a) it makes parameters sensitive to the context, and b) it makes the evaluation of attributions relative to a context of assessment, rather than the context of the subject of attribution. Both aspects are connected to the idea that understanding attribution concerns the satisfaction of certain epistemic *tasks*.

The basic motivation for introducing the notion of an epistemic task comes from certain features of the way in which we discriminate epistemic success. Successful epistemic states are successful because they arrive as types of actions are performed, which have epistemic value. Here I will call the types of actions whose performances have epistemic value as epistemic

[23] DeRose (2005) argues that contextualists about 'knowledge', unlike invariantists, can also accommodate cases like this where the standards used for attribution depend on a broader range of contextual factors.

[24] Technically, we could have different threshold functions for different context types, but I don't see the need for this complication here.

[25] Another way to put this is to have a C function that picks out a task-context from a pair of contexts, which in turn is passed to a P funtion that picks the appropriate evaluation function P for that context.

tasks.[26] The performances of epistemic tasks have value because, when satisfied, they characteristically change our epistemic state to better serve our needs. The resolution of inquiries, by solving a problem or getting an answer to a question is the prototype of an epistemic task in the sense intended here: it is epistemically valuable because it signifies our pass from a state of ignorance to a state of knowledge (assuming, of course, that this is the goal of inquiry). Consequently, epistemic tasks constitute epistemic goals (and specific functions for epistemic states). This suggests that epistemic tasks determine the context of epistemic evaluation: whether some state is successful or not depends on the task at hand in the sense that some ongoing task (a task-context) fixes the conditions for success. In the case of understanding, there will be tasks that involve the change of the part of the subject-at-hand's epistemic state that is relevant to what we talk about when we talk about their understanding–call these *understanding-tasks*.[27] When we make judgments about someone's understanding, we are making judgments about how they do in tasks of that sort.

The parameters that are relevant for epistemic evaluation need to fit the tasks at hand. We should worry, then, that by keeping the parameters of evaluation invariant, we could not accommodate ways in which the tasks vary across contexts–that by not allowing them to vary, they can no longer fit the tasks. Consequently, the parameters of evaluation should be taken to be context-sensitive.[28] In the terminology of the preceding section, the point is that the context selects P functions with different parameters in different contexts.

To illustrate how this is intended to work, and to compare it with other models, consider the following scenario. Suppose that in context A accuracy of judgment is required for understanding attributions, whereas in context B accuracy and coherence are required. Parameter-contextualism would describe this as a case where for context A a function P_a parametrized on accuracy is used, whereas in context B a function $P_{a,c}$ parametrized on accuracy and coherence is used (context selects the relevant P function). The parameter-invariantist alternative to this is to have an enriched invariant evaluation function that can accept all possible parameters. In the example, then, the parameter-invariantist would have both contexts use a function $P_{a,c}$, parametrized on both accuracy and coherence. To accommodate the relative

[26] A terminological note: we often will talk as if the tasks are the performances themselves. By giving you an order, you are given a task to do. By following that order, you engage in the task. You cannot engage in the *type* of action, but in *tokens* of that type. Nevertheless, the task you are given is not just the specific token-action you engage in (in a sense), but also, and perhaps primarily, the action-type (presumably, the token-action does not exist until you engage in it).

[27] Note that these are not specifically tasks *to understand* something. In general, many epistemic tasks will affect one's understanding-state, including one's pursuit of inquiry and reflection. I take it as a given that understanding as a personal-level concept is really about how a person's epistemic state integrates.

[28] McKinnon (2015) offers an account that can be presented along these lines for the norms of assertion. Her view is that what counts as warranted assertion is the possession of adequate supportive reasons, but that what counts as adequate supportive reasons varies with the context. One way to elaborate this is that having supportive reasons (i.e., being assertible) is such that for every context, there is a function P selected by the context which maps statements to values of appropriateness of assertion, such that for an statement to be assertible, it must satisfy P.

Bas Van Fraasen's (1988) pragmatic theory of explanation also seems to follow this model. According to him, explanations should be understood as answers to why-questions. Since according to him the content of why-questions, is context-sensitive, in this theory whether a sentence answers a why-question (that is, whether it is an explanation) depends on the context.

importance of parameters in different contexts, the parameter-invariantist should have something like a context-sensitive system of weights for the parameters.[29]

The parameter-invariantist proposal must assume that there is a definite (and finite) set of possible parameters–otherwise, plausibly, there would not be a definite evaluation function.[30] Furthermore, it is necessary for the view that all possible parameters could be evaluated at once by the same catch-all function.[31] In contrast, the parameter-contextualist view only requires the existence of ad-hoc evaluation functions, and these can be as simple or complex (in terms of the range of parameters that they take) as needed.

The advantageous simplicity of parameter-contextualist models becomes clear if we consider what we need to say if we want to filter out parameters as irrelevant in certain contexts. At first sight, it might appear as if invariantist accounts cannot accommodate cases of this sort–after all, the invariantist assumes that all parameters are evaluated at all contexts. But this underestimates the potential moves available to the invariantist. The real reason why we should adopt the parameter-contextualist model is the cost of the moves that invariantists *do* have available.

What are those moves? Some strategies are very similar to those available in the case of attributions of knowledge. For example, it could be argued that matters of parameter-relevance belong to the pragmatics of attributions, as a concern about their assertibility. For example, if in some pragmatic context parameter A is salient, attributions of understanding to states which are deficient according to that measure would end up unassertible, since asserting them would wrongly implicate that the salient parameter is satisfied non-deficiently.[32] If we evaluate the resulting models algorithmically, however, pushing filtering to the pragmatics will be wasteful of work at the level of semantic processing. Remember that the function that parameter invariantists need to remain fixed in all contexts evaluates all possible parameters each time. However, *ex hypothesi*, not all parameters will be relevant in each case–evaluating them is unnecessary. The parameter-contextualist model allows for a more efficient picture of the required processing, since everything can be evaluated on-demand, at the cost of some processing up front for the selection of relevant parameters.

---

[29] A different invariantist alternative would be to have a single parameter P function that was parametrized on the whole of the relevant context. Adopting this alternative would merely push all the work in the model into the internal structure of the evaluation function–which could yield a model that is trivially equivalent to the contextualist one. This seems to me like a merely aesthetic choice instead of a serious alternative.

[30] I think that this is problematic over the background of the idea that understanding attributions should themselves be understood in the context of epistemic tasks. Epistemic tasks emerge in the context of reflective and, more importantly, *extensible* practices–the agents engaged in them should worry whether there are novel ways for them to achieve their goals. While surely there are limits to the extensibility of epistemic practices, I find it doubtful that we can judge from within them that no further expansion is possible in the relevant senses. Epistemologically, we are not in a position to know the limits of the extensibility of these practices.

[31] One could interpret this as the idea that the agent that is evaluated has a state that has definite values for all the evaluated parameters, and that they *can* be evaluated as the function describes–call this the 'abstract' version of the evaluation function. However, we want this evaluation to be in effect 'plugged' into the larger procedure that ultimately yields attributions and judgments, so it is more natural to think that the function stands for an real procedure. If we had a separate function that, out of the abstract function, selected a subset of values that corresponded to a more limited set of parameters, we would have something that in effect is equivalent to what the parameter-contextualist model is intended to describe. But then the invariantist's objection would turn into a terminological squabble.[fn:squabble]

[32] Contrast with the pragmatic accounts of knowledge attributions in Rysiew (2001) and Brown (2005).

A different strategy is to concede some degree of context-sensitivity at the level of the semantics, but to argue that this happens within the evaluation function itself, rather than externally. The idea here would be that, as part of the catchall evaluation function, the evaluation mechanism would take the context as an input, pick out the salient features that need to be evaluated or filtered out, and then, proceed to assign epistemic values accordingly. Now, it is clear that this mechanism would be for all intents and purposes equivalent to the parameter-contextualist mechanism–indeed, depending on how the mechanism worked, it could be identical strictly speaking.[33]

The parameter-invariantist story cannot be *as* simple as the preceding discussion suggests, however. In the context of an understanding-tasks, agents will not only make outright attributions of understanding: they will also need to make comparative and evaluative judgments concerning understanding. The semantics of those statements will also rely on evaluative functions for epistemic states. It makes sense, then, that agents make use of the same evaluative functions for outright attributions and comparative and other evaluative judgments in the same task-contexts. But if the class of parameters that are needed to make comparative and evaluative judgments is wider than the class of parameters that are needed to make outright attributions, then *a fortiori* we have to assume that the evaluative functions at play are those that take the most parameters. The invariantist can push this point and ask: why not, then, accept that the evaluative function at play is indeed the catch-all evaluative function they propose? We can reply on two fronts. First, even if we need to assume that the relevant evaluative functions take more parameters than those that are minimally needed for outright attributions, this does not entail that they have to be the catchall function that parameter-invariantism requires, as the invariantist objection suggests. Second, we do not even need to assume that the same evaluative functions are used for both classes of statements (indeed, earlier I argued that we would want to have the semantics of outright attributions and comparative judgements relatively distinct). The idea that the same functions can be used for both cases is attractive because it helps in unifying the cases, which otherwise might appear somewhat disparate–even though the connection between both classes of statements is not straightforward, this does not necessarily mean that there are no systematic links between them. However, deploying the same function in both cases is not the only way to ensure this unity. Since tasks take time, an alternative is to have the evaluators apply different but similar functions at different points of their engagement with the tasks in response to prompts within the task-context. In engaging with an understanding-task, Smith has in mind some things that can serve as parameters for evaluation, but plausibly not everything that could be evaluated. Then, when asked to assess whether the tasks are being performed well or badly in certain respects, he adjusts how he evaluates them accordingly.

Endorsing the idea that understanding attributions are context-sensitive, in the way that parameter-contextualism suggests, opens up the possibility that the properties which have been taken as candidates for constituting the relevant success/attribution conditions (such as factivity, safety, well-connectedness, explanatory power, and what have you) are not constitutive of understanding success/states as such–rather, they might be characteristic of specific task-contexts. One advantage of this approach is that a large class of potential counterexamples to whatever substantive theories of understanding we have becomes less problematic, as long as it is possible to identify that these putative counterexamples rely on shifting or misidentifying the relevant epistemic-tasks. At the same time, this allows a task-based theory to explain why

[33] Cf. footnote [fn:squabble].

different criteria seem to be constitutive, or at the very least characteristic, of large classes of understanding states. In the task-based model, to different types of tasks will correspond different evaluation functions. Plausibly, then, similar tasks will converge into the same kinds of evaluation mechanisms. In other terms, mechanisms that serve similar functions in the context of specific epistemic tasks will have similar evaluation profiles.

For example, it is rather plausible that understanding attributions in the context of a large class of tasks will require the satisfaction of some factivity condition: many epistemic tasks will be truth oriented, or require that the agents that engage in them guide themselves by true judgements. For similar reasons, understanding attributions in the context of a subset of those tasks will require the satisfaction of some knowledge condition, as a stabilizing condition: assuming that knowledge is resistant to revision, it will ensure that the beliefs that one is guided by in engaging in the task are properly truth-oriented.[34] If we understand attribution in terms of the satisfaction of success conditions, then it makes sense that in many cases we will want to achieve successes that cannot be later invalidated–this is the kind of robustness that is required in practices where the function of understanding attributions might be to track good explainers among a population (cf. Hannon (2019)). All of this, however, is a contingent feature of the tasks that we are concerned with, instead of a constitutive requirements for the possession of understanding. Other tasks may not require the satisfaction of any of these conditions at all.

The worry now is whether it is appropriate at all to make outright attributions of understanding in contexts where any particular set of conditions are not met (for example, factivity or knowledge). The matter is complicated because,as I already pointed out, there may be a gap between the conditions for the possession of understanding and the conditions for outright attribution–remember my methodological worries in the introduction. One possibility here is that ordinary attributions are only made in the context of a proper subset of understanding tasks.

3.3 Relativity

By making the assessment context matter for the evaluation of epistemic states that yields the appropriateness of attributions (by giving, at least partially, the standards and parameters of evaluation), TBMU exhibits features of a relativist semantics, in the sense of MacFarlane (2014).

A consequence of this is that, in the context of a single epistemic-task, different agents could appropriately select divergent evaluation functions. Of course, this will depend on what is the epistemic-task at hand. For certain tasks, such as those which appear in scientific practices, there should not be more than one appropriately selected evaluation function, since it will be essential for those tasks that different agents can engage with them and converge in specific solutions. In those cases, if two agents come up with different solutions, at least one must be mistaken. If our model of attribution allowed that different evaluation functions were used appropriately in such cases, this could be used against the model. So we must make sure that the TBMU does not have this consequence. On the other hand, it must be possible for the model to allow divergence of evaluation functions in cases where tasks naturally allow it.

Since it is the fact that the TBMU model allows for divergence that makes it different from the more popular alternatives, we should take a look at the kind of situations that *do* call for this

[34] Other conditions could play stabilizing roles: for example, the possession of certain dispositional profiles such as competences, skill, or other epistemically virtuous characters.

feature before moving on to examining how the model avoids the pitfalls that I have just mentioned.

Whereas the epistemic tasks of, for example, science, concern subject matters for which there presumably are fixed facts which are entirely independent from perspectives, there are subject matters for which that is not the case.[35] A full defense of the point is beyond the scope of this paper, but I hope that at least some people would agree with me that *artworks* are an example of a kind of object (in a broad sense) for which these assumptions that enable convergence are mistaken. For my purposes here it suffices that I make the case that, for people who share this intuition, a model like TBMU should be attractive in a way that traditional models cannot be for lack of resources. Whereas for some objects such as natural systems, processes or even some artifacts (like machinery and instruments) one can think that optimal understanding will involve the possession of all the truths about the phenomena,[36] this does not seem adequate of the case of understanding art. Differing and even incompatible ways to think about artworks often seem–are?–viable. How to deal with this from the perspective of a theory of understanding attributions?

Here is one sketch of an answer. Some of the epistemic tasks involved in assessing artworks are not, like the activity of answering a question, the kind of activity that aims towards a specific end (answering a question). In many cases, we should think of the activities of assessment of artworks as open activities, that rather have their own continuation as a goal.[37] It is aesthetically interesting to try to come up with new views of old, well known works, that can suit better the peculiarity and idiosyncrasy of those who engage in the evaluation. From the perspective of activities of this sort, the properties of artworks are not fully determinate.[38] Thus, the nature of understanding *in the context of these tasks* should allow for the possibility of divergent but appropriate understandings. The potential lack of convergence is a feature, not a bug, of how these practices proceed. But even when there is a lack of convergence, this is also a contingent matter.[39]

The TBMU model allows for precisely this. The context of the assessors of artworks may include the requirement that they try to diverge in the content of their understandings, so the evaluation functions that they deploy will not necessarily be homogeneous–on purpose. However, because the evaluation functions are indexed to the task-context, which includes both the context of the atributee and the context of the assessor (in the model the evaluation function is doubly-indexed to capture this idea), the evaluation functions cannot be so disparate as to give rise to doubts that they are pertinent to the same kind of tasks. This is important because it helps avoid the objection, commonly raised against relativist approaches, that in the model 'everything goes'. The open-game metaphor I used before can help illustrate the point: even though the rules of an open-game can change, it nevertheless remains a game, and not something else.

[35] Even for the scientific case, some argue against the fixity thesis. See Giere (2006).
[36] Cf. Kelp (2015).
[37] I understand these activities as games, along the lines of Suits (2005).
[38] Elgin (2017, ch. 8) ascribes a view of this kind to Goodman (1978). She herself presents a more nuanced view concerning the function of disagreement between interpretations. However, it seems that even if her view does not assume that a converging understanding can result from developing those disagreements, in practical terms she suggests that it is valuable to incorporate knowledge of other perspectives into one's own understanding.
[39] Lack of convergence is not a reason to think that understanding of art objects is not grounded on cognitive states. Davies (2011, ch. 7) gives a rich account of musical understanding which is sensitive to the issue while adopting a moderately cognitivist stance.

What might cause worry for cases like these is what kind of story we have to adopt for comparative judgements. Since the evaluation functions shift, there may be cases where the evaluation of one subject A will differ in kind from the evaluation of some other subject B. As long as the task is the same for both, we can rely on a certain overlap of the parameters involved, but this does not mean by itself that comparisons are possible. In effect, in order to make comparisons for subjects who are being assessed by different evaluation functions, we need to adopt a different perspective from which to make those evaluations coherent. This process may be prone to failures, so that comparative judgements in these domains will not be reliable. Perhaps, like Elgin (2017) suggests, navigating through such comparisons and potential disagreements should be seen as a dialectical process: rather than arriving at the relevant contexts with a standard to measure understanding, we are forced to, together in dialogue, come up with criteria that all participants in the task at hand can assent to. This, by itself, might be the source of further divergence and open-endedness in our engagement in these tasks.

3.4 Subject matters and tasks

Ordinary understanding attributions ascribe their subjects with understanding *of* something or the other, either explicitly or by reference to something implicit in the conversational context (a dialogue: 'Does Mary understand Diophantine functions?' 'Yes, she does.'). A theory of attributions cannot fail to say something about this, but the point is particularly worrisome in the present context, since the issue is connected to the individuation of what I have called understanding-tasks.

Understanding some X may itself be type of epistemic task where understanding X to some standard is the goal of one's epistemic efforts. This is not always the case: it is possible for someone to come to understand by engaging in a task with a different goal–for example, by trying to find the solution to a mathematical problem, one may come to understand some mathematical idea related to the problem.[40] The problem of the individuation of tasks has different profiles in both cases.

In the first case, taking on the task fixes the criteria by which the task can be evaluated. A tempting idea is to think that setting oneself to understand X, one needs to have in mind what 'understanding X' amounts to. However, it is not right to think that what one has in mind when taking on the task is something precise. Rather, the right picture is that *during* the engagement with the task itself one settles into more specific criteria. One case where this is clear is where, by recognizing a lack of understanding of the target, one raises the standards for saying that one has achieved whatever epistemic goals one has. Engaging with some subject matters can result in felling like one has less understanding overall about a subject matter even though one has gained understanding about it. Remember that understanding may very well be evaluated across a multitude of dimensions–it can happen that as one's understanding increases in breadth, for example, more of one's understanding of specific parts of the target subject-matter is shallow. Or one can understand a specific point to a great degree of depth without one's understanding gaining any breadth–without, crucially, being able to place one's understanding of the subject matter into a true picture of the subject matter.

[40] It may even be the case that what one understands is not even related to the task at hand: during the process of trying to achieve one's epistemic aims one could spend some effort in trying something that does not contribute to the task at hand, but which nevertheless results in coming to understand something new.

All of this holds already for the first-person evaluation of understanding. In cases where the assessor differs from the understanding subject, there is also a question of how to identify the task and the criteria of evaluation that it brings with it. In some cases it will be possible for the assessor to simply ask the subject what he intends and by which criteria to evaluate–we can simplify these cases by reducing them to first-person cases if we assume that the communication is transparent. Of course, communication can be opaque, so in practice in many cases we will need an alternative mechanism for the assessor to identify the task and criteria. This is also needed to handle cases where it is not possible for the assessor to communicate with the subject. The material that the assessor works with in these cases is whatever trace he has of the activity of the subject, plus his understanding–as evaluated from his own context–of those activities. It is not possible to evaluate the understanding of someone whose activities are incomprehensible for us–outside of someone's activity, their epistemic state is opaque.[41]

This, however, is not easy. We can take a hint from Hasok Chang's (2009) idea that epistemic activities require certain ontological principles to be satisfied. Counting requires discreetness, narration requires subsistence, and so on. The satisfaction of these principles serves as a fixed parameter for the evaluation of those activities: counting something that is not discrete does not make sense, for example. This is what Chang means when he says that 'intelligibility is the performability of some epistemic activity' (p. 75). In this sense, here we are asking how to make sense of some presumed understanding task in its specificity: what parameters make sense to measure given the features of some subject's presumed tasks? How is understanding *itself* intelligible? My suggestion here will be that an assessor in need to reconstruct the context of a subject of understanding attribution has to build inductively some form of hypothesis about what the subject could have been doing in regards to the target of their understanding. Only then, with a guess about what the subject's task was, can the target project a way to evaluate the understanding of the subject. Chang claims that it is 'futile to tie understanding down to any particular activities' (p. 76). I agree in the sense that there is no single kind of task that gives rise to understanding. But I disagree if he also means that it is not possible to put together understanding and specific activities in a given context. To understand in a context just is to perform certain epistemic activities, and thus have a certain state, in a way that matches certain criteria.

In the second kind of case, where understanding evaluation comes in tasks with explicit goals other than understanding, the question of identifying the relevant understanding-task is again a mayor worry. One problem is whether it even makes sense to evaluate the understanding of subjects whose aims are not explicitly understanding-involving. Thus, for example, Chang rises the worry:

> Simply following an algorithm provides no relevant understanding to someone who is interested in some other epistemic activity, for example, visualizing what is going on, or giving a mechanical explanation. But for someone whose goal is to derive a prediction, there is surely the relevant sense of understanding in knowing how to apply the right tricks to derive the answer.

Sure enough, the subject in Chang's first scenario will not judge themselves as having gained understanding by following the algorithm. This can be simply explained by the fact that the task that they are engaged in does not afford the opportunity for the kind of evaluation of one's

[41] Also from the first-person: my own epistemic state is something that I also need to inquiry into.

epistemic state that could even in principle lead to first-person understanding attributions.[42] Could the subject not be assessed as having gained understanding if the context of the assessor demanded it or even allowed it? If so, that would give further support to my proposal that for attribution, the context of the subject is not enough.

What, then, is the mechanism by which assessors are supposed to evaluate the understanding of subjects who are engaged in tasks which are not explicitly aimed at understanding? Like in the cases where the tasks are explicitly aimed at understanding but where assessors cannot communicate with the subjects, assessors need to project a model of evaluation on whatever traces of the subject's activity they will have available. Since in this case there is no assumption that the subject aimed towards understanding, the assessor cannot make a guess about what the subject could have aimed to do, and what criteria they (the subject) would have evaluated their activity. One alternative is for the assessor to make additional assumptions about the epistemic state of the subject, and how the activities that they performed could have affected this state. The method could be useful for cases where the subject aimed at understanding, but it is not clear what parameters should count as relevant in the subject context.[43]

The inductive nature of the mechanisms by which assessors pick out the parameters by which to evaluate claims of understanding makes them prone to failure. It is no wonder, then, that in many cases there is room for disagreement about outright attribution and comparative judgements. On top of this, there is the issue that the target of understanding states may also be elusive in the way that tasks are. Consider the following scenario: Agatha aims to argue in favor of Platonism in mathematics, and to that end evaluates the Quine-Putnam indispensability argument. She assents to the premises, and thinks that the argument is valid. Now, assume that the argument cannot in fact support the conclusion. Did Agatha understand the argument? Let's assume, for the sake of the argument, that she did not–if she had understood the argument, she would not have assented to the conclusion. Nevertheless, if we narrow on other features of the case, perhaps we should say that Agatha did understand–not the argument, but certain features of the argument. And we could appeal to this fact to explain why some people would claim that, contrary to our supposition, Agatha did understand the argument. What has happened there is a subtle change of topic, which has led some assessors astray. These complications are unavoidable unless we pay attention to the issue of aboutness in understanding.[44] For now it will suffice to raise it.

4 TBMU and other accounts of attribution

Now that we have all the pieces of the TBMU model at hand, we can compare it with some alternative accounts from the literature. Since there is no shortage of accounts in offer, I will limit myself to three accounts that have aspects that are interesting to discuss here. Specifically, I

[42] Chang cautiously says that the subject's activity does not provide them with *relevant* understanding, leaving open the possibility of the subject acquiring understanding in a sense that is not relevant for the subjects's task. Nevertheless, he does not make anything out of that possibility.

[43] A different strategy, which builds on this idea, could be for the assessor to evaluate the state of the subject as if they had acted aiming towards understanding. While it may be attractive in some cases, this runs the risk of distorting the features of the epistemic activities that the subject did explicitly engage in. One context in which the strategy might play a more useful role could be in cases where the assessor has evidence that the subject did aim at understanding in some sense.

[44] Cf. Yablo (2014).

will my compare proposal with those of Kelp (2015, 2017, 2018), Baumberger (2019; 2017), De Regt (2017) and Wilkenfeld (2018, 2013). It has to be noted that my worry here is purely the model of attribution (the TBMU is not a theory of understanding by itself), so I am not concerned with their substantive accounts of understanding apart from the way in which those interact with their models of attribution.

4.1 Kelp's idealism: well connected knowledge

Kelp's account (2017, 2015, 2018) is interesting to discuss because it is one of the earlier explicit attempts at answering the questions that we started with to an appropriate degree of precision, and it has various features that make it illustrative of how to account for some kinds of context-sensitivity.[45] Kelp's account of understanding is knowledge-first friendly: according to him, understanding phenomena is a form of well-connected knowledge. More importantly for our purposes, he assumes that there is, for any phenomenon, the possibility of *maximally well connected knowledge* (*mwck*)–this is the ideal of understanding (hence, his theory is 'idealist', using a turn of phrase from Khalifa (2013)).[46] Someone who has *mwck* has the highest degree of understanding. Then, we can construct all other degrees of understanding as a measure of the distance between the epistemic state of agents and *mwck*. Context-sensitivity comes into the picture in the form of a contextually-fixed threshold for outright attribution:

Well Connected Knowledge

> 'A understands P' is true in context c if and only if A approximates fully comprehensive and maximally well-connected knowledge of P closely enough to be such that A would (be sufficiently likely to) successfully perform any task concerning P determined by c, if, in addition, A were to have the skills needed to do so and to exercise them in suitably favorable conditions.

Kelp (2015) made the qualification that the target of the view was just understanding of *phenomena*, but Kelp (2017) drops this restriction. Given this, we can classify Kelp's account as a form of Threshold Invariantism, since the kind of parameters that determine the evaluation of the epistemic state of the agents who are attributable with understanding is fixed across all contexts. This results from Kelp underlying assumptions about inquiry, which are not entirely explicit: either understanding is only relevant in the context of the epistemic-task of inquiry, or inquiry is taken to be a sort of overarching epistemic task that we are not in a position to avoid or suspend (cf. Kelp (2018)). This is not the place to settle these questions, but it should not be hard to grant that there is space for divergent views on the matter. In any case, these points are connected with the issue, that Baumberger (2019) rightly points out, that there is no context-sensitivity in the model for what counts as ideal understanding.[47]

---

[45] In fact, my formulation of the problems of outright attribution and comparative judgements derive from his. Cf. Khalifa (2017, 4, fn. 1).

[46] Khalifa (2013, 2017) also applies this idealist strategy but offers a different account of what ideal understanding amounts to.

[47] It may be objected that I am being unfair with Kelp's account. In the model, the context also determines a set of potential tasks concerning the target of understanding. But this in fact is *not* an independent source of context-sensitivity. Rather, it is precisely what, from the context, fixes the threshold for the fixed-criteria that the model requires for attribution. Can't the the tasks that the subjects are expected to be able to successfully perform change with the context also?–if the context induced a restriction in the range of tasks, the model would become parameter-contextualist.

As it stands, then, I think Kelp's account overgeneralizes. On the other hand, the version that is restricted to phenomena, can be accommodated as a special case of the TBMU model: in the context of the task of inquiry, the kind of criteria that Kelp suggests could be selected for the evaluation of epistemic states. Within that context, however, there could be alternative schemes of evaluation against which KOU itself should be evaluated (for example, why prefer Kelp's account over a similar version that requires acceptance instead of knowledge?).[48]

4.2 De Regt: the context-sensitivity of intelligibility

Like in the current proposal, the context-sensitivity of understanding attributions is one of the main focuses of De Regt's (2017) proposal. We have to keep in mind that De Regt's theory is only meant to apply to scientific understanding, and more precisely to explanatory understanding - it does not intend to provide an account of understanding attributions of the generality of something like TBMU. Regardless, it supposes a model of attribution that has interesting features. The core idea of De Regt's proposal is that for an agent to have pragmatic understanding of a theory, the theory needs to be *intelligible* for them. Intelligibility is the value that agents attribute to the properties of a theory that facilitate its use (p. 40). The intelligibility of theories depends on the context: for example, a theory which requires heavy computation is contextually less intelligible that one that has lesser computational requirements, and in the extreme case, it might not be intelligible in the absence of the appropriate devices, abilities or opportunities. More fully, De Regt endorses

Criterion for Understanding Phenomena

> A phenomenon P is understood scientifically if and only if there is an explanation of P that is based on an intelligible theory T and conforms to the basic epistemic values of empirical adequacy and internal consistency.[49]

Here, it is not the presence of a contextually fixed parameter that explains the context-sensitivity of understanding, but the variability of the parameters that are relevant to intelligibility itself. The model of attribution is, consequently, a type of parameter contextualism (and thus as a precursor of a generalized model like the TBMU.) Different domains will specify different criteria of intelligibility (p.101), and within a domain, there might be a multitude of criteria for the intelligibility of theories (p. 102). So while the account includes a general criterion for the kind of parameters that are relevant to attributions (they must be relevant to the use of the target of understanding, namely, theories, and they must be constrained by concerns of adequacy and consistency),[50] it allows for a different kind of context-sensitivity than a model like Kelp's, against which it competes.

De Regt's theory allows two agents to possess the same theories and background knowledge of some phenomenon, and yet differ in terms of whether they understand it in the sense of the phenomenon being intelligible for them–a kind of subject-relativity that the TBMU also shares.

---

[48] The TBMU model is compatible with the substantive component of various accounts of attribution to be *optional* within the context of the task of inquiry. Consequently, the pull towards deciding which substantive model of evaluation is the correct one is minimized.

[49] It is important to note that De Regt's proposal does not talk about the conditions under which a subject understand some phenomenon, but the conditions under which a phenomenon is understood. De Regt is interested in the conditions under which a *community* can be attributed with understanding–he calls this the *meso-level* of scientific activity (cf. p. 90).

[50] Cf. De Regt's discussion of Van Fraasen's pragmatic theory of explanation, p. 127–128.

De Regt appeals to a distinction of levels in order to defuse the objection that this leads to a dangerous form of relativism: at the level of scientific communities, standards for evaluation of the intelligibility of theories and phenomena are rooted in shared practices, so that 'scientific understanding does not vary capriciously from one individual to the next'. In the TBMU, the worry is defused by the double context-relativity of attributions and the role of epistemic tasks. In effect, epistemic tasks in the sense of the TBMU, like De Regt's scientific communities, act at a level that is higher than individuals. In fact, it makes sense to see tasks as more fundamental: scientific communities exist because individuals engage in tasks collectively.

4.3 Wilkenfeld: multiple dimensions and compression

Another author who has paid attention to the context-sensitivity of judgements about understanding is Daniel Wilkenfeld (2013, 2015, 2018). In fact, Wilkenfeld has developed various accounts of understanding; here I will focus on the MUD account from Wilkenfeld (2015) and the compression account from Wilkenfeld (2018). In terms of the model of attribution, all of Wilkenfeld theories follow similar lines.

The core of the MUD account is that states of understanding can be evaluated along a multitude of dimensions, among which we find representational accuracy and intelligibility (Wilkenfeld suggests that there might be independent dimensions that cannot be captured in terms of these). In terms of attributions, Wilkenfeld accepts the kind of context-sensitivity that a model like De Regt's captures, at least in the sense that what intelligibility means depends on the context. Furthermore, he claims that the weight of intelligibility vs. representational accuracy is context-sensitive. Thus, in some contexts intelligibility will be the main criterion to assess understanding, whereas in others accuracy will matter more. Within the conceptual framework developed here, this can be captured by a weighted parameter contextualist model of evaluation. Properly speaking, Wilkenfeld (2015) does not offer an account of outright attribution. This is because he takes the semantics of outright attributions to derive from the semantics of comparative judgements, much like in Kelp's (2015) model (see section 2.1 above for the details of this strategy for giving outright attribution conditions).[51]

In Wilkenfeld (2018), the multiple dimensions theory is replaced by the idea that each context specifies a type of information that is relevant to whatever epistemic tasks the context presents, and that understanding agents in such context should be able to generate information of that kind. Thus, we get

Understanding as Compression

> A person p_1 understands object o in context C more than another person p_2 in C to the extent that p_1 has a representation/process pair that can generate more information of a kind that is useful in C about o (including at least some higher order information about which information is relevant in C) from an accurate, more minimal description length.

Presumably, then, there is a threshold fixed by the context for the degree that a person understands an object o (in that context) that would make outright attributions appropriate. For the kind of tasks that Wilkenfeld presumably is concerned with, the model is (except for the lack of double indexicality), roughly equivalent to the TBMU model proposed here; we can take it as a special case. It should be noted that, unlike with Kelp and De Regt's proposals, Wilkenfeld

[51] Cf. Wilkenfeld (2018), p. 6, fn. 16.

thinks that the compression account can generalize more broadly than as an account of scientific understanding. However, it needs to be shown that it can be accommodated to a much wider array of tasks in the sense of the TBMU.

4.4 Baumberger: taking degrees seriously

Baumberger (2019) offers a thorough discussion of some of the issues that someone facing the issue of understanding attributions and comparative judgements faces. Unlike here, where I have focused on the way that context-sensitivity is handled, Baumberger focuses on the way in which we can make sense of degrees of understanding. He recognizes three strategies to do so: a) *Minimalist*: One establishes a minimal condition for outright attribution, and then accounts for degrees in terms of ways in which one can improve on that minimal understanding state (he attributes this approach to Grimm (2014)), b) *Idealist*: One established a condition that represents maximal or ideal understanding, and then accounts for degrees in terms of the distance between states and the ideal state (as we just saw, this is the approach that Kelp and Khalifa take), and c) *Direct*: One gives a standalone account of degrees, and then builds the accounts for outright attributions and comparative judgements on top of that (he attributes this strategy to Wilkenfeld (2013), and it is his own choice).

Baumberger spends most of his time defending his own substantive account of the dimensions of objectual understanding, but he does explicitly attempt to go beyond the models of other authors to capture a broader range of kinds of context-sensitivity. It is necessary to point out that the target of his analysis is a notion of objectual understanding, which he construes as a ternary relation between a subject, a subject matter and a theory (the subject matter is understood by means of the theory). This restriction doesn't narrow the scope of the theory as much as in the case Kelp's account, but there may still be a worry about the generality of the view. Instead of semantic clauses for outright attributions and comparative judgements, Baumberger offers what he calls and explication of objectual understanding:

Objectual Understanding

> An epistemic agent A understands a subject matter S by means of a theory T only if A commits herself sufficiently to T of S, and to the degree that (1) a grasps T, (2) T answers to the facts, and (3) A's commitment to T of S is justified.

In this model, the gradability of understanding comes from four angles: there is a relevant degree of commitment to T, a degree of grasp of T, a degree of how much T answers to the facts, and a degree of how justified the commitment to T is. Importantly, for the case of comparative judgements only the last three are taken to matter for evaluative contexts, whereas the first dimension is constitutive: how committed the subject is to the relevant theory does not make his understanding better or worse–but for them to be said to understand, they have to be committed to it.

Baumberger accepts that *how* the four conditions need to be satisfied in order to make outright attributions true depends on the context. He says:

> The explication admits of ascribing minimal understanding to an agent who satisfies the commitment condition and one of the conditions (1)–(3) to some degree. However, there will hardly be a context in which such an understanding is good enough to ascribe outright understanding to the agent. Individually necessary and collectively

> sufficient conditions for outright understanding in a given context can be arrived at by defining a context-specific threshold for each of the four conditions, or by combining such a threshold for the commitment condition with one for conditions (1)–(3) together in order to allow for trade-offs between them.

Baumberger is not entirely definite on what he intends the mechanism of attribution to be–in fact, he suggests a pair of alternatives of how we can obtain the conditions for outright attribution at a context. But even those can be understood in two different ways, in terms of the categorization of models of attribution that I have laid out here.

The first is to have him endorse a form of Threshold Invariantism with a complex condition. In both cases, we can assume that there is an evaluation function that aggregates values for all four conditions. The differences lie in how thresholds are handled. In one of the options suggested, the context fixes a threshold specification for the four conditions–in the form of a tuple $\langle T_c, T_g, T_r, T_j \rangle$. A different way to understand Baumberger's proposal is to assume that the threshold specification is a tuple $\langle T_c, T_x \rangle$, where T_x is a value that comes from an aggregating function for the other three conditions. However, this option does not seem to offer much more other than introducing a different sense in which understanding would be gradable, besides the gradability of the underlying conditions separately. It could be that the aggregation of values is complex, as I illustrated in section 2.1 with my discussion of weighted models, but Baumberger does not elaborate on this point.[52]

The second option is to understand Baumberger's model as a kind of restricted Parameter Contextualism. In this construal, there is a set of possible dimensions of evaluation to understanding (grasp, rightness and justification), and a general constraint on states (commitment). The context fixes an evaluation function that aggregates the commitment constraint with a combination of the evaluative parameters, and a corresponding threshold specification for the appropriateness of outright attributions. This is similar to the kind of models we saw in section 2.3, where a condition on the available evaluation functions is introduced to ensure the appropriate connection to epistemic value (the rationale for the restriction is different, however). How to understand the threshold specifications presents the same questions as before, yielding two possible reconstructions.

Against the first way of understanding Baumberger's proposal, the proponent of TBMU can rehearse the arguments already given when the view was compared with an 'enriched parameters' version of invariantism. When understood along the lines of the second option, the view can be seen as a restriction of the TBMU model to a class of tasks (namely, understanding a

[52] Another way to understand what Baumberger suggests is that there is a context-insensitive threshold for the satisfaction of the commitment condition, and a context-sensitive threshold specification that aggregates the values of the other three conditions. If the threshold for the commitment condition wasn't fixed differently from the rest, there would be no need to single it out–put differently, the (non-structural) information that can be extracted from an n-tuple $\langle P_1 \ldots P_n \rangle$ is the same as the information that can be extracted from the structure $\langle P_1, \langle P_2 \ldots P_n \rangle \rangle$. This construal does not cohere with Baumberger's explicit claim that commitment is also context-sensitive, but he does not show how the gradability of commitment plays a role in the assessment of outright understanding attributions. This is important because the degree of commitment is not an evaluative dimension of understanding, so it is not clear what function it may play in attribution besides as a simple requirement (contrast with Baumberger and Brun (2017), where commitment is also an evaluative parameter). Yet another option would be to take one threshold for commitment and another for all three of the other conditions. But this seems unappealing, since it would introduce an apparently arbitrary connection between those conditions.

subject matter through a theory) that require commitment constitutively, and where relativistic effects can be discounted.

5 Conclusion

To summarize: We have examined in detail how different models of understanding attributions can be developed, from simple invariantist ones to complex context-sensitive and relativist ones. I have argued that allowing for more thorough-going kinds of context-sensitivity can account for more patterns of understanding attribution. I have sketched a new kind of model of attribution (both parameter-contextualist and relativist), and defended it against some potential objections. I have also compared my approach with those of some recent works on understanding–and shown that my approach compares favorably. If nothing else, I hope to have clarified the kind of issues that arise when one tries to account for understanding attributions in detail, and the kind of modeling options that are available.

Acknowledgments

Special thanks to the anonymous reviewers from *Theoria*, and to the reviewers of a previous version of this paper elsewhere. Thanks to Jan Heylen, Lars Tump, Kristine Grigoryan, Giulia Lorenzi, and JiMin Kwon and the rest of the members of the Understanding Reading Group 2020–2021 for their suggestions throughout the stages of writing this paper.

References

Baumberger, Christoph. 2019. "Explicating Objectual Understanding: Taking Degrees Seriously." *Journal for General Philosophy of Science* 50: 367–88.

Baumberger, Christoph, and Georg Brun. 2017. "Dimensions of Objectual Understanding." In *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, edited by Cristoph Baumberger Stephen Grimm and Sabine Ammon, 165–89. Routledge.

Bengson, John. 2018. "The Unity of Understanding." In *Making Sense of the World: New Essays on the Philosophy of Understanding*, edited by Stephen Grimm, 14–53. Oxford University Press.

Brown, Jessica. 2005. "Adapt or Die: The Death of Invariantism?" *Philosophical Quarterly* 55 (219): 263–85. https://doi.org/10.1111/j.0031-8094.2005.00398.x.

Camp, Elisabeth. 2014. "Logical Concepts and Associative Characterizations." In *The Conceptual Mind: New Directions in the Study of Concepts*, edited by E. Margolis and S. Laurence, 591–621. MIT.

Carter, J. Adam. 2014. "Relativism, Knowledge and Understanding." *Episteme* 11 (1): 35–52.

Chang, Hasok. 2009. "Ontological Principles and the Intelligibility of Epistemic Activities." In *Understanding Scientific Understanding*, edited by Henk W. de Regt, 64–82. University of Pittsburgh Press.

Davies, Stephen. 2011. *Musical Understandings & Other Essays on the Philosophy of Music*. Oxford: Oxford University Press.

De Regt, Henk. 2009. "The Epistemic Value of Understanding." *Philosophy of Science* 76: 585–97.

De Regt, Henk. 2017. *Understanding Scientific Understanding*. Oxford: Oxford University Press.

De Regt, Henk, and Dennis Dieks. 2005. "A Contextual Approach to Scientific Understanding." *Synthese* 144: 137–70.

DeRose, Keith. 2005. "The Ordinary Language Basis for Contextualism." *The Philosophical Quarterly* 55 (219): 172–98.

Elgin, Catherine. 2017. *True Enough*. Cambridge, MA: The MIT Press.

Giere, R. 2006. *Scientific Perspectivism*. Chicago: University of Chicago Press.

Goodman, Nelson. 1978. *Ways of Worldmaking*. Indianapolis: Hackett.

Grimm, Stephen. 2014. "Understanding as Knowledge of Causes." In *Virtue Epistemology Naturalized: Bridges Between Virtue Epistemology and Philosophy of Science*, edited by Abrol Fairweather, 329–45. New York: Springer.

Hannon, Michael. 2019. *What's the Point of Knowledge?* London: Oxford University Press.

Hills, Alison. 2009. "Moral Testimony and Moral Epistemology." *Ethics* 120: 94–127.

Hills, Alison. 2016. "Understanding Why." *Nous* 50 (4): 661–88.

Hills, Alison. 2018. "Aesthetic Understanding." In *Making Sense of the World: New Essays on the Philosophy of Understanding*, edited by Stephen Grimm, 159–76. Oxford University Press.

Kelp, Christoph. 2015. "Understanding Phenomena." *Synthese* 192 (12): 3799–3816.

Kelp, Christoph. 2017. "Towards a Knowledge-Based Account of Understanding." In *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, edited by Cristoph Baumberger Stephen Grimm and Sabine Ammon, 251–71. Routledge.

Kelp, Christoph. 2018. "Inquiry, Knowledge and Understanding." *Synthese* 198: 1583–93.

Khalifa, Kareem. 2013. "The Role of Explanation in Understanding." *British Journal for the Philosophy of Science* 64 (1): 161–87.

Khalifa, Kareem. 2017. *Understanding, Explanation, and Scientific Knowledge*. London: Oxford University Press.

Kvanvig, Jonathan. 2009. "Responses to Critics." In *Epistemic Value*, edited by A. Haddock, A. Millar, and D. Pritchard, 348–61. Oxford University Press.

MacFarlane, John. 2009. "Nonindexical Contextualism." *Synthese* 166: 231–50.

MacFarlane, John. 2011. "Relativism and Knowledge Attributions." In *Routledge Companion to Epistemology*, edited by Sven Bernecker and Ducan Pritchard, 536–44. Routledge.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.

McKinnon, Rachel. 2015. *The Norms of Assertion: Truth, Lies, and Warrant*. London: Palgrave McMillan.

Moltmann, Friederike. 2009. "Degree Structure as Trope Structure: A Trope-Based Analysis of Positive and Comparative Adjectives." *Linguistics and Philosophy* 32 (1): 51–94. https://doi.org/10.1007/s10988-009-9054-5.

Pinkal, Manfred. 2005. "On the Logical Structure of Comparatives." In *Natural Language and Logic*. Vol. 459. Lecture Notes in Computer Science. Springer Verlag.

Pritchard, Duncan. 2009. "Knowledge, Understanding and Epistemic Value." *Royal Institute of Philosophy Supplement* 64: 19–43.

Rosenberg, J. F. 1981. "On Understanding the Difficulty in Understanding Understanding." In *Meaning and Understanding*, 29–43. Berlin: De Gruyter.

Rysiew, Patrick. 2001. "The Context-Sensitivity of Knowledge Attributions." *Noûs* 35 (4): 477–514. https://doi.org/10.1111/0029-4624.00349.

Schwarzschild, Roger, and Karina Wilkinson. 2002. "Quantifiers in Comparatives: A Semantics of Degree Based on Intervals." *Natural Language Semantics* 10 (1): 1–41. https://doi.org/10.1023/A:1015545424775.

Strevens, Michael. 2004. "Bayesian Confirmation: Inductive Logic, or Mere Inductive Framework?" *Synthese* 141 (3): 365–79.

Strevens, Michael. 2020. "Grasp."

Suits, Bernard. 2005. *The Grasshopper: Games, Life, and Utopia*. Edited by Thomas Hurka. Broadview Press.

Van Fraasen, Bas. 1988. "The Pragmatic Theory of Explanation." In *Theories of Explanation*, edited by Joseph C. Pitt. Oxford: Oxford University Press.

Weisberg. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Wilkenfeld, Daniel A. 2013. "Understanding as Representation Manipulability." *Synthese* 190 (6): 997–1016.

Wilkenfeld, Daniel A. 2015. "MUDdy Understanding." *Synthese*. https://doi.org/10.1007/s11229-015-0992-x.

Wilkenfeld, Daniel A. 2018. "Understanding and Compression." *Philosophical Studies*. https://doi.org/10.1007/s11098-018-1152-1.

Wilkenfeld, Daniel A., Dillon Plunkett, and Tania Lombrozo. 2016. "Depth and Deference: When and Why We Attribute Understanding." *Philosophical Studies* 173 (2): 373–93. https://doi.org/10.1007/s11098-015-0497-y.

Wilkenfeld, Daniel A., Dillon Plunkett, and Tania Lombrozo. 2018. “Folk Attributions of Understanding: Is There a Role for Epistemic Luck?” *Episteme* 15 (1): 24–49. https://doi.org/10.1017/epi.2016.38.

Yablo, Stephen. 2014. *Aboutness*. Princeton, NJ: Princeton University Press.