



Explicating Objectual Understanding: Taking Degrees Seriously

Christoph Baumberger^{1,2} 

Published online: 16 September 2019
© The Author(s) 2019

Abstract

The paper argues that an account of understanding should take the form of a Carnapian explication and acknowledge that understanding comes in degrees. An explication of objectual understanding is defended, which helps to make sense of the cognitive achievements and goals of science. The explication combines a necessary condition with three evaluative dimensions: an epistemic agent understands a subject matter by means of a theory only if the agent commits herself sufficiently to the theory of the subject matter, and to the degree that the agent grasps the theory (i.e., is able to make use of it), the theory answers to the facts and the agent's commitment to the theory is justified. The threshold for outright attributions of understanding is determined contextually. The explication has descriptive as well as normative facets and allows for the possibility of understanding by means of non-explanatory (e.g., purely classificatory) theories.

Keywords Understanding · Explanation · Prediction · Explication · Grasping · Justification · Commitment

1 Introduction

It is a commonplace that the sciences and the humanities aim at a better understanding of various aspects of the world. But until recently, philosophers have hardly considered understanding a cognitive achievement that needed to be addressed for its own sake. Epistemologists have been preoccupied with propositional knowledge, and philosophers of science have focussed on explanation instead of understanding. Understanding was either seen as a mere counterpart of explanation that does not call for an independent analysis, or as a psychological by-product of explanation that falls outside the domain of a philosophy only concerned with the rational reconstruction of science.

✉ Christoph Baumberger
christoph.baumberger@usys.ethz.ch

¹ Institute for Environmental Decisions, ETH Zurich, Universitätsstrasse 16, 8092 Zurich, Switzerland

² Centre for Philosophy of Natural and Social Science, London School of Economics, Houghton Street, London WC2A 2AE, UK

The recent turn to understanding is based on the conviction that a suitable notion of understanding can serve certain theoretical purposes that cannot be fulfilled by the concepts of knowledge and explanation alone. First, some philosophers argue that making sense of the cognitive achievements of science and of scientific progress requires an adequate account of understanding, because science aims not only to acquire isolated pieces of knowledge about the world but to understand it, and the idealized models underlying this understanding involve falsehoods that are incompatible with knowledge, which according to a standard analysis implies truth (Elgin 2004; de Regt 2015; Potochnik 2015; Dellsén 2016a). Second, there is a growing awareness that understanding can neither be identified with explanation nor reduced to a subjective feeling evoked by explanations (de Regt 2009). Still, understanding and explanation seem to be conceptually linked. Some claim that understanding is the goal of explanation (Grimm 2010) and that any account of explanation has to “tell us what kind of understanding scientific explanations provide and how they provide it” (Friedman 1974, 14). Others suggest reversing the traditional order of explication and analysing the concept of explanation in terms of a suitable notion of understanding (van Camp 2014; Wilkenfeld 2014). Third, epistemologists have found it difficult to account for a supposed special epistemic value of knowledge, not shared by states that fall short of knowledge, such as true belief. Some argue that such a value problem does not arise for understanding and that epistemology should therefore turn its attention to understanding (Kvanvig 2003; Pritchard 2010). Finally, there are strong intuitions that epistemic justification needs to be accessible to the agent and that it requires embedding individual beliefs into a coherent web of beliefs. It has been proven difficult to save these internalist and coherentist intuitions in an account of knowledge, but access to the reasons for belief and the ability to connect a belief with others seem to be central to understanding (Kvanvig 2003). Thus, a suitable conception of understanding may accommodate deep-seated intuitions about epistemic justification that have been rejected as intuitions about knowledge (Greco 2010).

The aim of this paper is to defend an explication of objectual understanding—understanding a subject matter by means of a theory or theory-like representation—that helps to make sense of the cognitive contributions and goals of science. “Science” is thereby understood in a very broad sense, familiar from the German term “Wissenschaft”, which covers not only the natural sciences, but also the formal sciences, the social sciences, and the humanities. My focus is on the first of the mentioned roles for a concept of understanding, but the suggested concept will also prove suitable for some of the other roles. I argue that an explication of objectual understanding needs four conditions. In a paradigmatic case, understanding a subject matter by means of a theory requires that the theory answers to the facts, the agent grasps the theory (i.e., is able to make use of it), commits herself to the theory, and her commitment to the theory is justified. With the exception of commitment, each condition constitutes an evaluative criterion for assessing how good the agent’s understanding is. The threshold for outright¹ attributions of understanding is determined contextually. Before I develop and defend the explication (Sects. 3 and 4), I motivate my approach by drawing some lessons from two deficits of the current debate about understanding (Sect. 2).

¹ Following Kelp (2015, 3805, fn. 6), I use “outright” for “not involving degrees”. An assertion of the form “S understands M”, e.g., is an outright attribution of understanding, in contrast to assertions of the form “S understands M to some degree” or “S₁ understands M better (or less good) than S₂”.

2 Constraints on an Account of Understanding

A diagnosis of two shortcomings of the current debate about the nature of understanding raises the following questions: What kind of account of understanding do we need and what are its criteria of adequacy? How should we deal with the fact that understanding comes in degrees? Answering these questions leads to some constraints on an account of understanding.

2.1 Two Shortcomings of the Current Debate

In recent years, several accounts of understanding have been developed.² The most prominent conditions suggested can be grouped into four types. A *factivity condition* captures the idea that understanding must be adequately tied to the facts, which is usually spelled out in terms of truth (Kvanvig 2003; Pritchard 2010). The idea underlying an *ability condition* is that someone who possesses understanding can do certain things, for example, apply it to actual and counterfactual cases by inferring conclusions and giving explanations (de Regt 2009; Hills 2016). A *belief condition* is based on the idea that a representation can only afford an agent with understanding if it is somehow hers (Kvanvig 2003; Pritchard 2010). What suggests a *justification condition* is that understanding is an achievement of an agent, which requires that the agent can provide good reasons for her beliefs (Elgin 2007; Kvanvig 2003; Pritchard 2010), or that her understanding results from a reliable process (Grimm 2017). In philosophy of science, “justification” is rarely used, but the idea is familiar, for example when it is claimed that the understanding-providing representation needs to meet logical and empirical requirements and additional theoretical virtues, such as simplicity and explanatory power (de Regt 2009).

All four conditions are contested. Some argue that understanding is not even moderately factive because it can be generated by idealized models and superseded theories (de Regt 2015; Elgin 2007). Others argue that the abilities a putative understanding enables are not relevant to its status (Strevens 2013), or can be explained in terms of knowledge (Kelp 2015; Khalifa 2012). It has even been argued that understanding neither requires justification nor belief, because it can be based on defeated evidence (Dellsén 2016b; Hills 2016; Wilkenfeld 2017). Needless to say, proponents of the suggested conditions have defended and sometimes modified them in different ways (see Baumberger et al. 2017a).

This debate is at risk of running idle, for at least two reasons. First, even though most authors agree that understanding comes in degrees and that it depends on the context how good someone’s understanding must be to ascribe her outright understanding, these insights have not been taken seriously enough. If they had, some disagreement about the necessary conditions for understanding would be less grave than it first seemed. If we agree that a suggested condition admits of degrees and constitutes an evaluative criterion, disagreement about how well it must be met in a certain context for outright understanding is comparatively undramatic. Whether someone can in some context be credited with understanding a subject matter becomes less urgent if we agree that her understanding would at least be minimal, for instance because the inferences she can draw about the subject matter are very limited, or because her theory is inconsistent with her background beliefs.

² The majority of accounts address “explanatory” understanding of why something is the case rather than objectual understanding, but this does not matter for my diagnosis.

Second and even more importantly, it is unclear how the dispute about the correct account of understanding can be resolved since it is unclear what exactly such an account aims at and thus what its criteria of adequacy are. Usually no such criteria are mentioned, and they differ widely, for example, depending on whether an account aims at a conceptual analysis in the traditional sense or at a Carnapian explication. I submit that if the aim is a concept of understanding that serves some of the theoretical purposes mentioned in the introduction, we should seek an explication rather than a conceptual analysis. In what follows, I argue this point with respect to the purpose of making sense of the cognitive contributions and goals of science, and discuss how an explication best deals with the fact that understanding comes in degrees.

2.2 What Kind of Account Do We Need?

A conceptual analysis in the traditional sense typically takes the form of a reporting definition that states individually necessary and jointly sufficient conditions. If such a definition has no formal flaws, it is adequate if definiendum and definiens are intensionally equivalent. The definition has thus exclusively to be tested against (intuitive) judgments about whether actual and hypothetical cases fall under the defined concept. Counterexamples leave us with no option but to revise the definition. Whether the definition serves any theoretical purpose is not a condition of adequacy and needs to be seen after the concept has been analysed. For my purpose, the prospect seems dim. An analysis which states only conditions that are indeed necessary is likely to define a notion of minimal understanding that is hardly suitable to make sense of the complex cognitive achievements of science. Moreover, its conditions are unlikely to be sufficient for ascriptions of understanding in demanding contexts. Giving up the idea of necessary and sufficient conditions and characterizing a family resemblance concept instead does not solve the problem because the adequacy of such a characterization is still independent of whether it serves our purpose. In everyday as well as in scientific contexts, the use of “understanding” is likely to be too diverse and too vague for an account of the cognitive contributions of science. But neither should we seek a stipulative definition of “understanding”, for example as the name of a cognitive goal of science whatever it turns out to be. Since the adequacy of such a definition is independent of whether the defined concept has any similarities with our ordinary concept, we might end up with an account of a cognitive goal that is not any more an account of understanding. The aim is rather to develop a concept of understanding that captures key features of our ordinary use of the term in relevant contexts, and is suitable for the purpose of making sense of the cognitive contributions of science. Developing concepts that are both similar enough with our ordinary concepts and theoretically useful is the basic idea of Carnap’s method of explication.³

The method of explication replaces an inexact concept (the explicandum) for some theoretical purpose with an explicitly introduced and at least as exact concept (the explicatum) by incorporating it into a target theory. In a first step, the explicandum must be clarified, which calls for disambiguating the explicandum-term, specifying its relevant use, logical form and range of meaningful applications. Moreover, it needs to be specified into which

³ The classical exposition is Carnap (1962, §§ 2–3); see Brun (2016) for a detailed discussion. Following Brun (2016, 1216–1217), I use “concept” to refer to a term together with rules for its use and leave it open how they are related to the abstract entities that Carnap (1962, 7–8) calls “concepts”.

target theory the explicatum is to be integrated and which purpose it should serve in this theory. In a second step, an explicatum must be introduced by specifying rules for its use in terms of the target theory. Ideally, this is done by a definition. The adequacy of the explicatum must be assessed in light of the theoretical purpose it is intended to serve in the target theory. Carnap specifies four criteria of adequacy that are all a matter of degree. First, explicatum and explicandum need to be sufficiently similar. This requires neither intensional nor extensional equivalence, but only that the explicatum can be used in place of the explicandum in relevant contexts. The similarity criterion aims to ensure that the explication does not change the subject. The next three criteria concern the theoretical usefulness of the explicatum. The explicatum should be sufficiently exact, as fruitful as possible (i.e., allow us to formulate many laws or generalizations featuring the explicatum), and if two explications both meet these criteria satisfactorily, simplicity (of the rules for using the explicatum and of the generalizations that include it) can be used to select one.

The philosophical practice of testing conditions of application against (intuitive) judgments about actual and hypothetical cases has its place in the method of explication. It comes into play when we assess whether the similarity criterion is sufficiently met. But the role of this practice is restricted since similarity is compatible with reclassification of pre-theoretically clear cases and only one among several criteria of adequacy. Each “counterexample” gives us a choice whether we should reject or accept it as a case of application of the explicatum and revise our explication. The decision is guided by considerations about the fulfilment of the criteria of adequacy with respect to the purpose the explicatum should serve: Does rejecting the example make the explicatum too dissimilar to the explicandum? Does accepting the example reduce the fruitfulness or exactness of the explicatum? Since I am looking for a concept of understanding that helps to make sense of science, it can for instance be reasonable to reject as counterexamples cases in which we are inclined to ascribe understanding to higher animals even though they do not (and cannot) meet all of the suggested conditions.

2.3 How to Deal with the Fact that Understanding Comes in Degrees?

Even though everyone accepts that understanding comes in degrees, most accounts address outright understanding. However, three strategies to account for degrees of understanding can be distinguished; each has their proponents in the current debate (Baumberger et al. 2017a, 26–27).

The first strategy explains degrees of understanding on the basis of an account of outright understanding. An account of minimal understanding is developed that specifies the conditions for outright attributions of understanding. Degrees of understanding are then introduced by pointing out how minimal understanding can be improved. Grimm (2017), for example, assumes a minimal understanding that is achievable for young children and higher animals and consists in reliably grasping dependency relations, and then shows how this “animal” understanding can be improved to become “reflective”. A second strategy starts with an account of maximal understanding and explains degrees of understanding in terms of approximations to maximal understanding. This strategy can be combined with a contextualist semantics for outright attributions of understanding, which determines how close one must come to maximal understanding of a phenomenon to be truly attributed outright understanding in a given context. Khalifa and Kelp both pursue this strategy but disagree with respect to the standard for maximal understanding. Khalifa (2013a) identifies it with knowing the best explanation of a phenomenon, Kelp (2015) with having fully

comprehensive and maximally well-connected knowledge of the phenomenon. A third strategy is to directly explicate what it means for an agent to understand something to a certain degree. Wilkenfeld (2015), for example, suggests that the extent to which someone understands a phenomenon depends on how accurate her representation of the phenomenon is, and on the kinds of good inferences she can make about the phenomenon. The threshold for outright attributions of understanding can again be determined contextually.

The first strategy faces a dilemma. If the specified conditions for outright understanding do not comprise all (and maybe only very few) dimensions with respect to which understanding can vary, the explication does not provide a sound basis for an account of degrees of understanding. As a result, it becomes unclear why good understanding is not a combination of understanding with something else. Grimm's (2017) account of animal understanding risks being speared by this horn of the dilemma, since reflective understanding is not simply a more reliable grasp of more dependency relations, but additionally requires the ability to articulate the dependencies. On the other hand, if the specified conditions for outright understanding do comprise all evaluative dimensions, then either the explication is likely to be too demanding, or it becomes an instance of the third strategy. The explication is likely to be too demanding if in every context all conditions need to be met for outright understanding; it turns into an instance of the third strategy if it allows in some contexts to ascribe understanding even though not all conditions are met.

The second and the third strategy have better prospects of avoiding this dilemma. Since they do not start with an account of minimal understanding, they are in a better position to specify all evaluative dimensions. And they are less prone to set the bar for outright understanding too high because they do not impose a non-contextually determined minimal threshold. One might think that starting with an account of maximal understanding is best suited to specify all evaluative dimensions. But the second strategy does not force their proponents to specify the maximum in a way that is helpful for an account of degrees of understanding. Khalifa characterizes maximal understanding in terms of whatever science takes to be the best explanation of a phenomenon and withholds from explicitly specifying evaluative dimensions; he holds that "descriptions of scientific practice are rich enough to give us anything else we could want from the epistemology of understanding" (Khalifa 2017, 142). Kelp distinguishes only two dimensions, namely how comprehensive and how well connected someone's knowledge about a phenomenon is, which may be seen to measure breadth and depth of understanding respectively (Kelp 2015, 3812).⁴

However, the main reason why I prefer the third strategy over the second is that it is not committed to the idea of a non-contextually determined maximal understanding. The claim that maximal understanding is context-independent can either be construed as the claim that it is independent from any context, or as the claim that in every context, the maximum is of the same type (e.g., knowing the best explanation). It is unclear whether the first interpretation makes sense and whether there is a specification of the second claim that proves true. For example, if some phenomena do not admit of an explanation but can nonetheless be understood, as Kvanvig (2009) argues, then maximal understanding of these phenomena cannot consist in knowing the best explanation of them. One might object that such phenomena do not admit of maximal understanding. But it seems plausible that maximal

⁴ Kelp and Khalifa could specify sub-dimensions, e.g., via conditions for knowledge. They probably have not developed their accounts in such a direction because their primary motivation for pursuing the second strategy is to explicate understanding in terms of knowledge rather than to provide a theory of degrees of understanding or the kinds of factors such a theory would have to take into account.

understanding of a phenomenon is the best understanding the phenomenon admits. Kelp will object that maximal understanding should rather be identified with fully comprehensive and maximally well-connected knowledge. But even if Kelp's account seems more promising, it may still be preferable to provide an explanation of degrees of understanding that does not presuppose a non-contextually determined maximal understanding.

3 Explication of Objectual Understanding

Before I can introduce the explicatum, the explicandum must be clarified, and the target theory and the theoretical role of the explicatum in this theory need to be further specified.

3.1 Identification of the Task

Examples of what I call “objectual understanding” are understanding climate change through climate models, the tides in terms of Newton's theory of gravitation, social segregation with the help of an agent-based model, or World War I through a historical work. In all these examples, we seek more than a single explanation of one phenomenon, and strive to understand a subject matter by means of a theory or theory-like representation.

This holistic sort of understanding first contrasts with what is sometimes called “atomistic understanding” (Pritchard 2010, 14), namely “explanatory” understanding why something is the case by means of an explanation (e.g., why the global mean surface temperature has increased rapidly in the last decades), and other forms of “interrogative” understanding, such as understanding what something is by means of an explication (e.g., what climate change is). Different types of interrogative understanding are typically part of objectual understanding. Understanding what counts as climate change and why the global mean surface temperature has increased, for example, are partially constitutive of understanding climate change. Second, “objectual understanding” is sometimes used in the literature in the wider sense of all forms of understanding that take a noun phrase as their object, including understanding of other things than subject matters (e.g., someone's action), and by other means than theories or theory-like representations (e.g., maps). Finally, understanding by means of a theory has to be distinguished from understanding a theory (e.g., understanding climate change vs. understanding a climate model) (Baumberger and Brun 2017, 166–167).

Objectual understanding is a relation between an agent, a subject matter and a theory. All three relata should be conceived in a very wide sense. In everyday attributions of understanding, the agent is typically an individual person. But recent work on cognition has stressed how crucial the social division of epistemic labour and the use of external devices such as instruments and material representations are in our cognitive practices. Since the phenomena of “distributive” and “extended” cognition are particularly important within science, an explication of understanding should allow the agent to be a collective agent (e.g., a scientific group or community) and/or an extended system (e.g., a system consisting of a scientist or a group of scientists and a computer) (Ylikoski 2014, 328). The notion of a subject matter in my characterization of objectual understanding is intended to cover objects of understanding that might also be described as a field (e.g., genetics), a topic (e.g., justice) or a system (e.g., a historical inheritance system). The notion of a theory is intended to apply not only to scientific theories in a narrow sense, but also to non-empirical theories and theory-like representations such as narratives (e.g., history works),

mathematical models (e.g., numerical climate models) and conceptual models (e.g., diagrammatic climate models). Furthermore, I do not presuppose a specific account of the nature of theories, and allow that theories can include, and sometimes even largely consist of, non-propositional representations, such as diagrams and graphs.

The target theory in which the explicatum should serve its role is an epistemological theory of different cognitive contributions and goals of science and different forms of scientific progress. This theory is not simply given; the explication is rather one step in developing such a theory. The theoretical purpose of the explicatum within this theory consists in making sense of understanding as one important cognitive contribution and goal of science. More specifically, the explicatum should account for the fact that understanding comes in degrees, allow us to acknowledge that paradigmatic scientific theories admit at least some understanding of their subject matters, and do justice to the fact that scientific understanding usually involves idealized models.

3.2 Introduction of the Explicatum

An example of good understanding will help to determine the various dimensions of objective understanding. This is in line with my argument that an explication of understanding should neither start with minimal nor with maximal understanding, but rather directly explicate what it means to have some degree of understanding.

Climate scientists claim to have a reasonably good (although imperfect and incomplete) understanding of climate change by means of complex climate models, which consist of mathematical equations that are based on physical laws (e.g., of fluid dynamics) and empirical assumptions (e.g., about cloud albedo) (IPCC 2013). Defending such a claim for a particular scientist and model requires arguing several points (Baumberger and Brun 2017, 167–168).⁵ First, the model needs to adequately reflect the facts. This is compatible with the model being an idealized representation, and only requires that it represents sufficiently well those climate processes that make a difference to the essential character of the system. A Global Circulation Model that represents well a wide range of atmospheric, oceanic, sea ice and land processes can provide a good understanding of climate change even if it leaves out many climate features (e.g., biogeochemical processes) and distorts in different ways processes and aspects represented to make them mathematically and computationally tractable. Most obvious are idealizations in the case of parameterizations which account for the effect of processes (such as cloud formation) that are important for the model results but occur on a smaller scale than can be resolved in the model, or are too complex or not well-enough understood to be modelled in detail.

Second, the scientist does not only need to know about the model, but must grasp it in the sense of being able to use the model to answer a wide variety of questions about climate change. For example, how will different climate characteristics evolve if greenhouse gas concentrations increase, stabilize or decrease in particular ways? To what extent is the global warming of the last 50 years due to human causes? Does climate change affect extreme summer precipitation in the Northern hemisphere? Since the model equations are analytically intractable, the scientist must be able to run the model on a computer to approximate solutions numerically on a spatial and temporal grid for specific initial and

⁵ The example could be modified so that the agent is a group of scientists and the means by which the understanding is achieved a family of models.

boundary conditions, and to interpret the model results to provide projections (i.e., predictions relative to emission scenarios) and explanations that answer the target questions. Being able to efficiently use a model in this way requires some qualitative comprehension of how the model behaviour emerges from the interaction of different model components (physical principles, parameterizations, initial and boundary conditions) and how it would change if some components were different in various ways.

Third, the scientist need not only be able but also disposed to use the model when a question arises concerning climate change which requires a model-based answer. This does not require that she believes the model, but that she accepts it as being suitable for the projections and explanations that answer the questions at issue.

Finally, the scientist's acceptance of the model needs to be the result of a reliable model evaluation and thus based on reasons which show that the model is sufficiently good in various respects. These include explanatory power, precision, empirical accuracy (i.e., fit of results to available data), robustness (i.e., fit of results to results of other models that differ, e.g., in number and complexity of processes included, parameterizations, parameter values, spatiotemporal resolution and numerical methods), and support by background theories and assumptions (e.g., support of model equations by physical theories from which they are derived and support of parameterizations by high-resolution models that explicitly resolve the involved processes).

This example suggests that an explication of objectual understanding needs four conditions to be able to account for good understanding: A rightness condition ensures that the theory sufficiently answers to the facts; a grasping condition that the agent is sufficiently able to make use of the theory; a commitment condition that she accepts the theory, for example for predictive and explanatory purposes; and a justification condition that the agent's commitment is based on reasons which sufficiently speak in favour of the theory. The four conditions are not specific to the natural sciences. Defending a student's claim that a historical work provides her with a good understanding of World War I requires to show that the work does justice to the historical facts, that the student grasps and accepts the work and is thus able and also disposed to use it to answer a variety of questions about World War I, including questions not explicitly addressed in the work, for example about what would have been the case if certain historical facts were different. Moreover, the student's acceptance of the work needs to be based on a reliable evaluation which sufficiently speaks in favour of the work and shows that it is sufficiently good with respect to accommodating historical evidence, explanatory power, intuitive plausibility, and support by background assumptions and theories, for example from psychology and sociology.

All four conditions admit of degrees, but only three of them constitute evaluative dimensions. How good an agent's understanding is depends on how well she grasps the theory, how well the theory answers to the facts, and how well her commitment to the theory is justified. However, the quality of an agent's understanding does not depend on how strongly she commits herself to the theory. Commitment comes in degrees, but the degree of understanding does not depend on the degree of commitment. This motivates the following explication⁶:

⁶ Cf. Baumberger and Brun (2017, 168–169), where all four conditions are treated as evaluative criteria. The explication does not feature an anti-luck condition since epistemic luck is compatible with some degree of understanding (Baumberger 2011; Hills 2016), which is even acknowledged by proponents of knowledge-based accounts of understanding (e.g., Khalifa 2013b; Kelp 2017). But epistemic luck diminish the degree of understanding. I account for this by building a reliability requirement into the justification criterion (see Sect. 4.3).

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,⁷ and
- (3) *A*'s commitment to *T* of *S* is justified.

It depends on the context how well the four conditions need to be met for an outright attribution of understanding to come out true. 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.

The explication is in line with Wilkenfeld's recent suggestion that "the quality of a state of understanding is evaluable along multiple orthogonal dimensions" (Wilkenfeld 2015, 4). But while Wilkenfeld argues for two such dimensions, namely intelligibility and representational accuracy which correspond to (1) and (2), my explication adds justification as a third dimension and commitment as a necessary condition. Moreover, grasping, rightness, justification and commitment correspond to the four most prominent conditions for understanding suggested in the current debate (see Sect. 2.1). But some conditions are construed in a more general way (e.g., commitment does not require belief but only acceptance, and rightness does not require factivity but is compatible with central falsehoods), and all conditions are taken as a matter of degree and, with the exception of commitment, as evaluative criteria. Taking grasping, rightness and justification as good-making features of understanding rather than as necessary conditions allows to account for the insight that these dimensions play an important role in the ascription and assessment of understanding without denying that there may be contexts in which we rightly ascribe understanding even though some of the conditions are hardly met or not met at all.

The above explication outlines the factors that a full theory of degrees of understanding would have to take into account without providing such a theory. If such a theory is possible at all, it will hardly enable us to determine in a quasi-mechanical way how good someone's understanding is and when it is good enough for attributing outright understanding in a given context. Determining this is subject to informal evaluation and judgement, for several reasons (familiar from Kuhn's 1977 discussion of theory choice): The conditions involve terms that are open to interpretation and call for further explications; there is no effective procedure for assessing how well the conditions are met; and it is usually impossible to specify in an exact and context-independent way how to weigh the different criteria, which trade-offs between them are admissible and what is required for them to be met to a sufficient degree. In a context in which we want to understand regional climate change in a way that allows reliable projections, an increase in similarity between model and target outweighs a decreased ability to grasp the model; in many educational contexts,

⁷ We might add "or get things right in some other way" to leave room for the idea that normative theories can be objective even if there are no normative facts they can answer to.

a good grasp of a model is more important than a high degree of similarity between model and target. In what follows, I focus on the individual conditions and leave the question of how to weigh them in a given context for another occasion.⁸

4 The Four Conditions

All four conditions require detailed discussion, which is outside the scope of this article. In this section, I provide preliminary characterizations and guidelines for a more thorough account. Any such account should meet three desiderata: It should acknowledge that the conditions admit of degrees (graduality), allow for idealizations (idealization), and apply to all kinds of theories and theory-like representations (diversity). Moreover, I defend the claim that rightness, grasping and justification are evaluative dimensions and that commitment is a necessary condition for understanding. My strategy will be to show that even those who reject rightness, grasping and justification as necessary conditions can and should accept them as evaluative dimensions.

4.1 Rightness

The mentioned desiderata cannot be met by an account of rightness that construes the theory-world relation in terms of the model-theoretic relation of isomorphism, as suggested by proponents of the semantic view of theories (e.g., van Fraassen 1980). Isomorphism accounts fail to meet the graduality and the idealization desiderata since models are either isomorphic to their targets, or they are not. Partial isomorphism accounts (e.g., da Costa and French 2003) do better. They divide a model structure into different substructures, and claim that a model is partially isomorphic to its target when a substructure of the model is isomorphic to a substructure of the target. Partial isomorphism comes in degrees, depending on the relative size of the isomorphic and non-isomorphic substructures, and accounts at least for some idealizations. However, neither isomorphism nor partial isomorphism accounts meet the diversity desideratum. They compare the mathematical structure of a model to the mathematical structure of a representation of the target and thus only apply to mathematical models and quantitative features (Weisberg 2013, 137–142). An account in terms of truth (e.g., Kvanvig 2003) may fare better since all kinds of theories can be evaluated with respect to their truth. Moreover, a truth-based approach may account for how

⁸ One of the reviewers pointed out that as long as we are lacking a story for weighing criteria (1)–(3), there is a wide range of relevant cases for which the explication does not tell us whether or not the explication applies. Think of cases in which one agent A satisfies any one of criteria (1)–(3) to a higher degree than some other agent B, whilst B satisfies some other of the criteria to a higher degree than A. If for many of these cases the explicandum admits of a clear verdict, then it is unclear whether the explication satisfies the similarity condition. I agree that more needs to be said about how to weigh criteria (1)–(3) in different contexts and how these assessments can be transferred to new cases. However, even without such a story, we can weigh the criteria in a given context by relying on intuitive judgments about comparative degrees of understanding. One might object that it is circular to use such judgments at the same time to decide how to weigh the criteria of the explication and to argue that the explication satisfies the similarity condition. This is like using the same data to calibrate and to test a model. But as we can split data and use one part of a dataset to calibrate a model and the other part to test it, we can rely on some intuitive judgments to decide how to weigh the criteria in a given context and on other intuitive judgments (e.g. about paradigmatic cases which differ exclusively with respect to one single criterion) to assess how well the similarity condition is satisfied.

well a theory answers to the facts by considering how many propositions of the theory are true, how central they are, and how close the false propositions come to truths. But even if such an account can be developed (which requires, e.g., a theory of truth approximation), it will not apply to non-propositional representations if their content is not fully explicable in terms of propositions and thus fail to meet the diversity desideratum.

More promising is an account of answering-the-facts in terms of similarity. Since anything is similar to anything else in some respect, similarity needs to be restricted. The rough idea is that a theory answers to the facts constituting the target system to the degree that the system as depicted by the theory is in relevant respects similar to the target system. Which respects are relevant depends on the context and the purpose that the theory is intended to serve (Giere 2010). In many contexts, the degree of relevant similarity is a matter of how many features (processes, factors) that make a difference to the behaviour of the target system are represented, and how well (detailed, precise, comprehensive) they are represented. Relevant similarity meets the desiderata because it comes in degrees, can be used to compare idealized models, and can relate all kinds of theories and theory-like representations to their targets (Weisberg 2013, 143). To develop these ideas, one can draw on similarity accounts within the literature on scientific representation.⁹ While most approaches say little about what similarity amounts to, how it depends on context and purpose, and how it can be assessed, such questions are addressed by Weisberg's (2013, Ch. 8) similarity account of the model-world relationship, which is based on Tversky's (1977) influential contrast account of similarity. The account starts with the idea that a "model is similar to its target [...] when it shares certain highly valued features, doesn't have many highly valued features missing, and when the target doesn't have many significant features that the model lacks" (Weisberg 2013, 144–145). This idea is transformed into an account of the model-world relation by considering in detail where the feature sets and the weighting function come from.

Most authors accept some kind of rightness condition for understanding, but even those who reject rightness as a necessary condition could acknowledge it as an evaluative dimension. Even if radically false idealized models and superseded theories can provide some understanding (de Regt 2015), the quality of the understanding can, *ceteris paribus*, be a function of how well the theory answers to the facts. Rightness is indeed an evaluative dimension of understanding since there are cases such that A_2 's understanding of target system S by means of theory T_2 is better than A_1 's understanding of S by means of T_1 but the only relevant difference is that T_2 answers better to the facts than T_1 , where A_2 can be A_1 at a later stage and T_2 a successor of T_1 .¹⁰ De-idealizations are a case in point.

⁹ In the literature on scientific representation, it is not always clear whether similarity is invoked to explain what makes something a representation, or what makes a representation adequate. Suárez (2003) has shown that similarity fails as an answer to the first question, but I am concerned with one aspect of the second question (the adequacy of a representation depends, e.g., also on how intelligible it is; see de Regt 2009).

¹⁰ Wilkenfeld (2015) argues against de Regt (2015) that an account which rejects rightness ("representational accuracy") as an evaluative criterion and accepts only the ability to use a theory (its "intelligibility") as such a criterion leads to implausible judgments of comparative degrees of understanding. Such an account implies, e.g., that a Ptolemaic astronomer of the sixteenth century understood the motion of the planets better than her Copernican contemporary who correctly puts the sun at the centre of the solar system, because the Ptolemaic theory is more intelligible to her and has a greater predictive power at that time. While Wilkenfeld takes this result to be counterintuitive, de Regt and Gijsbers (2017) are ready to bite the bullet. However, the decisive question is whose understanding was better at the fictitious moment in which both theories were equally intelligible and predictively successful. It seems clear that in this case, the Copernican astronomer understood the motion of the planets better than her Ptolemaic contemporary.

In climate science, two types are particularly important. The first concerns the scope of climate processes represented in the model and can be illustrated with the transition from Global Circulation Models (GCMs) to Earth System Models (ESMs). While GCMs represent already a wide range of processes of the physical climate system, ESMs additionally represent biogeochemical processes, such as the global carbon and sulphur cycles. This allows ESMs to explicitly simulate feedbacks between the changing physical climate and biogeochemical cycles that determine the greenhouse gas concentrations, which are in GCMs, for the most part, simply prescribed (Flato 2011). The second type of de-idealization concerns the level of detail and comprehensiveness with which processes are represented in a model. An example is the replacement of empirical parameterizations by sub-models that explicitly resolve the involved processes. Other things being equal, both types of de-idealization improve our understanding of climate change, but neither of them need lead to better predictions and retrodictions, at least not in the near term. A more faithful representation of cloud processes, for example, that performs better when tested individually may even lead to a poorer performance of the model as a whole if the model is biased with respect to aerosol concentration or humidity (Baumberger et al. 2017b). Assuming the performance is still reasonably good, the understanding is improved. The example suggests that this improvement is due to the increased similarity between model and target rather than to other factors, such as an improved ability to retrodict past and predict future climate.

The claim that the goodness of someone's understanding is a function of how well her theory answers to the facts is a *ceteris paribus* claim. Often, the *ceteris paribus* clause is not fulfilled since the increase in similarity between theory and target goes along with a decreased ability to use the theory or a decreased justification. Climate models that represent an ever increasing range of processes and explicitly resolve processes that in predecessor models are included via parameterization may become too complex to understand and too computationally intensive to be run on available computers. And models that represent (e.g., certain biogeochemical) processes that are difficult to observe or have not been systematically observed over a long period of time or over large spatial scales may be less justified than simpler models (Flato 2011, 783).

4.2 Grasping

A good starting point for an account of the grasping condition is to look at how scientists determine whether someone understands a theory. Exams play an important role here. We take exams to show whether a student has merely memorized a theory and maybe even memorized certain applications of the theory, or whether she is able to apply the theory to new examples. This suggests that knowing a theory by testimony and knowing in this way that a theory applies to certain cases is not the same as truly understanding the theory, which requires being able to make use of it. This, in turn, suggests that an agent grasps a theory to the degree to which she is able to apply it to actual and counterfactual situations (de Regt 2009; Stuart 2016; Newman 2017).

What the ability to apply a theory involves depends on the kind of theory at issue. In typical cases, it involves the ability to provide predictions and explanations in terms of the theory, given certain information about the target. However, a merely classificatory theory may give some understanding of the facts about a domain, even if it does not enable us to explain these facts. Gijsbers (2013) argues that classifying animals on the basis of their anatomical features allowed biologists of the eighteenth century to make correct predictions,

which is a cognitive achievement that deserves to be called understanding, but it did not allow them to explain the features they were able to predict. Khalifa (2012) objects that for every case of understanding without explanation, there exists a correct explanation that would provide greater understanding. But even if Khalifa were right,¹¹ this would not show that a general account of grasping should require explanatory abilities. Even if the ability to explain turns out to be necessary for good understanding, there may be lower degrees of understanding that can do without explanation. An account of grasping that meets the diversity desideratum should not even require predictive abilities, at least not in the literal sense of predictions in which they concern future events. The historical natural sciences do not make predictions, the formal sciences and many theories in the humanities not even retrodictions, but theories from these disciplines certainly provide understanding of their subject matters. However, the ability to use a theory always implies the ability to draw consequences of the theory about aspects of the subject matter, which is often not a matter of straightforward deduction, but involves plausible reasoning.

The ability to apply mathematical models involves the ability to accurately calculate results for different initial conditions, boundary conditions and parameter values, and to interpret the results to solve quantitative problems about the target. If the model equations can be solved analytically, the calculations may be performed with the help of pencil and paper or even mentally. In case of models consisting of analytically intractable equations, determining numerical values requires the ability to run the model on a computer to estimate solutions numerically. However, it is often possible to cheat one's way to a mathematical solution without really understanding it and the underlying model. A good grasp of a model requires therefore also the ability to solve qualitative problems by drawing consequences of the model without performing exact calculations (de Regt and Dieks 2005, 151). This requires a good comprehension of how the model behaviour emerges from the interaction of model components and how it would change if some components were different in various ways. This is why most novices who are almost as competent as their expert teachers with solving quantitative problems fail when it comes to qualitative problems (Newman 2017, 579). The ability to solve quantitative problems by calculating and interpreting results and the ability to solve qualitative problems by estimating consequences may in the case of non-mathematical (qualitative) theories correspond to the ability to answer questions about the target by explicitly going through the argumentation suggested by the theory, and the ability to estimate answers by drawing characteristic consequences of the theory without complete logical argumentation (cf. de Regt and Dieks 2005, 167, fn. 7). Moreover, a good grasp of a theory may involve the ability to assess the conditions and limits of its application, and to judge what types of results it allows about the target. This is particularly important in case of idealized models, the grasping of which requires some awareness of how they diverge from reality and of the conditions under which the divergences are negligible so that the models can be applied to the target.

The ability to use a theory comes in degrees, but at least in certain contexts, one can have some understanding even if one is unable to independently apply the theory to its target. I may have some understanding of global warming by means of a simple Energy-Balance Model, without being able to construct explanations and projections of temperature trends in terms of the model. It suffices that I can follow such explanations and projections when given by someone else. This requires more than the ability to repeat them, and more

¹¹ Khalifa's argument is directed against cases of understanding-why without explanation, which have been suggested by Lipton (2009), but it seems likely that it applies also to Gijsbers's example.

than a shallow semantic understanding of the model and its applications. Some grasp is needed of how the elements of the model are related and how it is that the model explains the explananda and entails the projections. This grasp involves the ability to provide a qualitative summary of the model (e.g., with the help of a diagram) and reformulate its applications in one's own words. The suggestion then is that a good grasp of a theory requires being able to apply it to its target (problem-solving abilities), and a minimal grasp being able to follow and reformulate such applications when given by someone else (comprehension abilities).¹² An account along these lines obviously meets the desiderata of graduality, idealization and diversity.

Most authors accept some kind of grasping condition for understanding and many tried to spell it out in terms of abilities, but even those who reject grasping-as-abilities as a necessary condition should accept it as an evaluative dimension. Even if some understanding is possible if the agent is unable to apply the theory and maybe even to reformulate such applications in her own words, the quality of her understanding is, *ceteris paribus*, a function of how well she grasps the theory. Suppose two climate scientists understand climate change in terms of the same climate model, their commitment to the model is equally justified and both are able to run the model on a computer and to interpret the results in order to answer a variety of questions about climate change. However, while the second scientist can, due to her experience in experimenting with models, qualitatively assess how the model results would change if certain model components were different, the first scientist is hardly able to anticipate any such counterfactual consequences without running the model on a computer. It seems clear that the understanding of the second scientist is better than that of the first, and that this is due to her better grasp of the model. But as in case of rightness, the *ceteris paribus* clause will often not be fulfilled. How well one grasps a theory may vary with the degree of one's justification, and the better grasp may depend on the fact that the theory is more idealized and thus less similar to the target.

4.3 Justification

Standard epistemological accounts conceive of epistemic justification as exclusively truth-conducive: justification speaks in favour of a belief being true or not being false. Since these accounts address knowledge and belief, they are not directly applicable to understanding and theories. They cannot even be adapted for an account of the justification condition for objectual understanding since they fail to meet the idealization desideratum. Idealized models and our commitment to them can be justified but we know that the models are not true,¹³ which requires that their justification is not only related to truth. This is in line with the claim, well known from philosophy of science, that the epistemic evaluation of theories needs to appeal to a plurality of epistemic goals—such as Kuhn's (1977)

¹² The terms “problem-solving abilities” and “comprehension abilities” are due to Newman (2012). One might argue that these abilities do not constitute the grasping but are grounded in it. But even if cognitive abilities are grounded in psychological states, this does not imply that the grasping involved in understanding should be identified with being in the underlying state. Identifying the grasping with having the abilities has the advantage that the resulting account of understanding allows for collective and for extended understanding.

¹³ For some idealizations, the justification may be primarily pragmatic, e.g., when the idealization only serves to make a model more computationally tractable. But this is not the case with minimalist idealizations which include only those causal factors that make a difference to the essential character of their target (Weisberg 2013, 98–103).

accuracy, consistency, broad scope, simplicity and fruitfulness—that are not exclusively truth-conducive and admit of trade-offs.¹⁴ This claim should be accepted if epistemic evaluation is understood as evaluation with respect to a theory's contribution to understanding (Baumberger and Brun 2017, 169–171): Increasing the explanatory power or simplicity of a theory can enhance our understanding of its target even if the resulting theory is less empirically accurate. A climate model that explicitly resolves cloud processes that predecessor models included via parameterization enhances our understanding of climate change due to its increased explanatory power (it can, e.g., explain feedbacks involving clouds), even if the model is empirically less accurate because of biases with respect to aerosol concentrations.

Theory-choice approaches meet the idealization desideratum, but they cannot simply be used as an account of the justification condition since they fail to meet the diversity desideratum. They are tailored to scientific theories and do not directly apply to non-empirical and normative theories. But they provide a good starting point for such an account. If accuracy is understood in a more general way and an internalist and a reliabilist requirement is added, they allow us to distinguish five dimensions of justification that are relevant in the context of objectual understanding: The degree to which an agent's commitment to a theory is justified depends (a) on whether the theory is internally consistent and on the degree of its external coherence with background theories and assumptions; (b) the degree to which the theory accommodates the available evidence, which includes observational or observation-based data in case of empirical theories, and intuitions in case of non-empirical theories (cf. Bealer 1996); (c) the degree to which the theory does justice to further epistemic goals, including generally relevant virtues (e.g., precision, simplicity, fruitfulness, broad scope, explanatory power, and completeness with respect to the subject matter) and virtues that are specifically relevant to certain kinds of theories (e.g., visualizability and causality); (d) the degree to which the agent is able to assess how well (a)–(c) are met; and (e) on the reliability of the theory evaluation.

Dimensions (a) and (b) assess how well the agent's theory answers to the facts. Some epistemic goals in (c) may contribute to such an assessment too; for example, broadening the scope of a theory may be a means of minimizing error since a theory with more (and more diverse) areas of application admits of additional tests (Douglas 2013). However, the epistemic goals in (c) primarily have other functions. They are used to assess the systematicity of a theory (Baumberger and Brun 2017), its intelligibility for the agent, that is, the ease with which the agent can make use of the theory (de Regt and Dieks 2005), and its relevance for specific problems (Hirsch Hadorn and Baumberger 2019). Which goals are relevant, how much weight they should be given and which trade-offs are acceptable depends on the subject matter and on the purpose the theory is intended to serve. If a climate model is developed for the purpose of understanding the basic mechanisms of global climate change, the model should be as simple as possible, but we may not insist on it being useful to effectively calculate exact figures for the key climate characteristics it involves. If, on the other hand, we want to understand regional climate change in a way that makes reliable projections available, we will require that the model be as detailed as necessary and of high

¹⁴ See Douglas (2013) for an overview. Some authors identify epistemic justification with truth-conducive justification and epistemic goals (or values) with truth-conducive goals (e.g., consistency and accuracy), which they distinguish from cognitive goals (e.g., broad scope, simplicity, and fruitfulness) that are valued for other reasons (Laudan 2004). I opt for a broader conception of epistemic justification and epistemic goals.

enough resolution for effectively computing the relevant climate characteristics with sufficient precision even if this means that the model gets incredibly complicated. Dimension (d) assesses the agent's metacognitive perspective on her epistemic situation, and (e) how unlikely it is that the evaluation process leads the agent to commit herself to a theory that does not sufficiently answer to the facts or is not relevant for the problem at issue. This last dimension accommodates the intuition that epistemic luck diminishes the degree of understanding (cf. Khalifa 2013b).

An account of justification that involves the dimensions (a)–(e) meets all three desiderata: it acknowledges that justification comes in degrees, accounts for idealizations since justification is not only related to truth, and applies to all kinds of theories.¹⁵

Most authors accept some kind of justification condition for understanding, but even those who reject justification as a necessary condition should acknowledge it as an evaluative dimension. Even if an agent can have some understanding by means of a theory that is not justified for her because her evidence is insufficient (Dellsén 2016a) or even defeated (Dellsén 2016b; Wilkenfeld 2017), the quality of the understanding is, *ceteris paribus*, a function of how well her commitment to the theory is justified. Suppose two climate scientists understand climate change by means of the same climate model, and both grasp the model equally well (i.e., are equally able to use it to address quantitative and qualitative problems about climate change). However, while the first scientist has largely to rely on the testimony of her peers for knowing how well the model meets some of the justification conditions, the second can assess by herself how empirically accurate and robust the model results are, how well the model coheres with her background theories and how well it performs with respect to epistemic goals such as simplicity, explanatory power and completeness. It seems clear that the understanding of the second scientist is better than that of the first, and that this is due to her better performance with respect to the internalist requirement. Or suppose both scientists are equally able to assess how well the model meets the justification conditions (a)–(c) but only the second scientist is able to rule out some incorrect rival model that could easily have met the rightness condition. Again, the understanding of the second scientist seems better than that of the first, and this is because his model evaluation is more reliable.

The result becomes even clearer if we assume that the model coheres better with the second scientist's background theories than with those of the first, for example because the model is incompatible with a background theory of the first scientists, or because the second scientist has a much broader range of background theories that support the model. Incompatibility with a background theory seriously reduces or even destroys one's understanding, and while it may be possible to gain some understanding by means of an "isolated" theory which is consistent with but mostly logically independent of relevant background theories, further support by background theories boosts understanding by integrating it into a wider picture (Baumberger and Brun 2017, 176). If we modify the example so that the two scientists grasp different climate models that equally well answer to the facts, their justification can differ (also) because one of the models performs better with respect to epistemic goals such as simplicity or explanatory power. Other things being equal, simplifying a model or improving its explanatory power will often advance the understanding of its target.

¹⁵ In Baumberger and Brun (2017), we develop a reflective-equilibrium account of the justification condition for objectual understanding, which involves dimensions (a)–(d).

As in case of rightness and grasping, the *ceteris paribus* clause will often not be fulfilled. How well an agent is justified in a theory varies with how well the theory answers to the facts and often also with how well the agent grasps the theory. Moreover, some dimensions of justification may not constitute evaluative sub-dimensions of understanding. In the above example, the justification of the models could also differ in that one of them is better supported by data, for example because the other represents processes for which only very limited data are available. But it is less clear whether such a difference influences the degree of understanding.

4.4 Commitment

Commitment as it figures in the explication of objectual understanding is an attitude towards some content. This attitude should not be identified with belief as taking a proposition to be true. An account of commitment in terms of belief meets the graduality desideratum but fails to meet the idealization and the diversity desiderata. Belief comes in degrees, but we do not believe idealized models we know to be false, and non-propositional representations are not suitable objects of belief since they are not truth-apt. Moreover, it is often impossible to explicate the content of such representations in terms of beliefs attributable to the agent.

Commitment should rather be explicated in terms of a broader notion of epistemic acceptance. An influential distinction between belief and acceptance is due to L.J. Cohen.¹⁶ According to him, to *believe* that *p* is to be disposed normally to feel it true that *p* (and false that non-*p*) when one is attending to issues raised by *p* or items referred by *p*. But to *accept* that *p* is “to treat it as given that *p*”, that is, “to adopt a policy of [...] including [*p*] among one’s premises for deciding what to do or think in a particular context, whether or not one feels it to be true that *p*” (Cohen 1992, 4). The notion of acceptance needs to be construed in a specific way to provide a sound basis for an account of commitment in the context of objectual understanding, which leads to further differences between belief and acceptance. First, belief is a propositional attitude, but the objects of acceptance include non-propositional contents and all kinds of theories and theory-like representations. Second, belief exclusively aims at truth, while acceptance relates to a plurality of epistemic goals. A climate scientist can accept a highly idealized climate model that deviates from truth if the deviations are negligible or compensated by the performance of the model with respect to epistemic goals such as simplicity and fruitfulness. Whether the deviations are negligible or compensated depends on the purpose the model should serve. Thus, third, in contrast to belief, epistemic acceptance is relative to epistemic purposes, which makes acceptance context-dependent in a way that belief is not. A scientist can accept a climate model to project the global mean surface temperature increase by 2100, but reject using the model to project changes in precipitation pattern in the Mediterranean area between 2050 and 2100.

Thus, an agent epistemically accepts a theory to the extent that she takes the theory to be useful for specific epistemic purposes, such as prediction, retrodiction and explanation, ideally based on an evaluation of the theory’s performance with respect to various

¹⁶ Cohen’s proposal has attracted some attention in the understanding literature, see, e.g., Elgin (2004), Potochnik (2015) and Dellsén (2016b). For alternative distinctions between belief and acceptance, see, e.g., van Fraassen (1980), Bratman (1992), and Elliott and Willmes (2013).

epistemic goals. Accepting a theory in this sense can be controlled voluntarily, while one may not be able to believe at will. An account of commitment in terms of acceptance meets all three desiderata: commitment comes in degrees,¹⁷ an agent can commit herself to an idealized model she does not take to be true, and to non-propositional and even non-verbal representations, such as diagrams and graphs.

Commitment is a necessary condition rather than an evaluative dimension. The examples in Sect. 3.2 provide some reason to assume that objectual understanding implies commitment. However, two examples do not establish a general claim, and they concern good understanding while it might be that lower degrees of understanding can do without commitment. More convincing would be examples in which we are inclined to withhold any understanding because the agent does not sufficiently commit herself to the theory in question. But in such examples, the agent typically also lacks justification and/or the rightness condition is not met.

Since it is difficult to provide a case that establishes commitment as a necessary condition, let us see whether there are any arguments against a commitment condition. I am not aware of such an argument, but Wilkenfeld (2017) and Dellsén (2016b) argue that understanding does not require belief. Do their arguments apply to commitment? Certainly not in case of Dellsén who aims to show that understanding may be accompanied by mere acceptance rather than by belief. Wilkenfeld (2017, 321–322), however, provides an example that can be interpreted as suggesting that understanding does not even require acceptance¹⁸: Richard, an established scientist, develops a detailed model of the explosion of the *Challenger* space shuttle, based on the idea that O-Ring failure caused the explosion. His model is entirely correct, but before he goes public, Richard is subjected to a deliberate cover-up with the result that his subjective credences indicate that he puts the probability for explanations in terms of his model at about 30% and the probability of their negation at about 70%. Richard does not believe his model, but he does not even seem to accept his model since he refrains from using it in explanations of the *Challenger* explosion. This case provides a counter-example to the necessity of a commitment condition only if it is plausible that Richard understands the *Challenger* explosion in terms of his model. However, it seems more natural to describe the case as one in which the explosion is understandable in terms of Richard's model, not by Richard but by those who have not been subjected to the same cover-up and are thus disposed to use the model in explanations of the explosion. We should distinguish between "subject matter S is understandable in terms of theory T" and "an epistemic agent A understands S in terms of T" as two different explananda. While an explication of the first does not require an acceptance or commitment condition, the second does.

5 Conclusion

I defended an explication of objectual understanding which combines a necessary condition with three evaluative dimensions: An agent understands a subject matter by means of a theory only if she commits herself sufficiently to the theory, and to the degree that she grasps the theory, the theory answers to the facts, and her commitment to the theory

¹⁷ This distinguishes commitment from acceptance in Cohen's sense, which does not admit of degrees.

¹⁸ Wilkenfeld's original example concerns explanatory rather than objectual understanding.

is justified. The explication has a descriptive and a normative facet (see Sect. 2.2). On the one hand, the similarity criterion requires that the explicatum conforms to a reasonable extent to the pre-theoretic, vague use of “understanding” as applied to subject matters.¹⁹ On the other hand, the explicatum can deviate from the pre-theoretic use in order to comply as well as possible with the criteria of theoretical usefulness (exactness, fruitfulness and simplicity). The adequacy of the explicatum must be assessed in light of the theoretical purpose of making sense of an important cognitive achievement and goal of science. My discussion of the climate modelling example provides some reason to assume that the suggested concept of objectual understanding serves this purpose. Further support calls for case studies, which I leave for another occasion.

Let me conclude by some advantages of the suggested explication. First, it shifts the focus from necessary conditions for outright attributions of understanding to evaluative dimension and comparative degrees of understanding. As Kelp (2017, 266–267) notes, this is a step forward since judgments of comparative degree are often less complex and more reliable than outright judgments, which is also the case with respect to understanding. Second, the suggested notion of objectual understanding provides a home for internalist and coherentist intuitions about justification without denying that in certain contexts someone can have some understanding without having an internalist and/or coherentist justification. Third, the explication allows for the possibility of understanding by means of non-explanatory (e.g., purely classificatory) theories. At the same time it acknowledges that explanatory power is typically an epistemic goal with respect to which the vehicle of understanding is evaluated. In many contexts, the ability to provide adequate explanations will even be required for ascribing outright understanding.

Finally, the suggested explication identifies understanding neither with a set of abilities, nor with a mental state. Accounts that take understanding to be a set of abilities (e.g., Ylikoski 2014; de Regt 2015) tend to have difficulties to make understanding answerable to the facts (van Camp 2014); accounts that construe understanding as a mental state tend to make it impossible to account for collective and extended understanding (Ylikoski 2014). The suggested explication makes understanding answerable to the facts and allows for the possibility that the epistemic subject is a collective agent (e.g., a group or a community) and/or an extended system (e.g., a person or a group together with external, material devices such as computers). This is secured by analysing commitment in terms of adopting a certain policy rather than as a mental state, and grasping in terms of abilities rather than as a mental act.²⁰

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

¹⁹ Assessing whether the similarity criterion is sufficiently satisfied may require some information about how to weigh the criteria of the explication in different contexts (cf. fn. 8).

²⁰ I would like to thank Claus Beisbart, Eric Brandstedt, Georg Brun, Finnur Dellsén, Victor Gijssbers, Gertrude Hirsch Hadorn, Mike Stuart and two anonymous reviewers for their helpful comments on earlier drafts.

References

- Baumberger, C. (2011). Types of understanding: Their nature and their relation to knowledge. *Conceptus*, 40(98), 67–88.
- Baumberger, C., Beisbart, C., & Brun, G. (2017a). What is understanding? An overview of recent debates in epistemology and philosophy of science. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 1–34). New York: Routledge.
- Baumberger, C., & Brun, G. (2017). Dimensions of objectual understanding. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 165–189). New York: Routledge.
- Baumberger, C., Knutti, R., & Hirsch Hadorn, G. (2017b). Building confidence in climate model projections: An analysis of arguments from fit. *WIREs Climate Change*, 8(3), e454. <https://doi.org/10.1002/wcc.454>.
- Bealer, G. (1996). “A priori” knowledge and the scope of philosophy. *Philosophical Studies*, 81(2/3), 121–142.
- Bratman, M. (1992). Practical reasoning and acceptance in a context. *Mind*, 101(401), 1–15.
- Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, 81(6), 1211–1241.
- Carnap, R. (1962). *Logical foundations of probability* (2nd ed.). Chicago: University of Chicago Press.
- Cohen, L. J. (1992). *An essay on belief and acceptance*. Oxford: Clarendon Press.
- da Costa, N. C. A., & French, S. (2003). *Science and partial truth*. Oxford: Oxford University Press.
- De Regt, H. W. (2009). The epistemic value of understanding. *Philosophy of Science*, 76(5), 585–597.
- De Regt, H. W. (2015). Scientific understanding: Truth or dare? *Synthese*, 192(12), 3781–3797.
- De Regt, H. W., & Dieks, D. (2005). A contextual approach to scientific understanding. *Synthese*, 144(1), 137–170.
- De Regt, H. W., & Gijsbers, V. (2017). How false theories can provide genuine understanding. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 50–75). New York: Routledge.
- Dellsén, F. (2016a). Scientific progress: Knowledge versus understanding. *Studies in History and Philosophy of Science Part A*, 56, 72–83.
- Dellsén, F. (2016b). Understanding without justification or belief. *Ratio*, 30(3), 239–254. <https://doi.org/10.1111/rati.12134>.
- Douglas, H. (2013). The value of cognitive values. *Philosophy of Science*, 80(5), 796–806.
- Elgin, C. Z. (2004). True enough. *Philosophical Issues*, 14(1), 113–131.
- Elgin, C. Z. (2007). Understanding and the facts. *Philosophical Studies*, 132(1), 33–42.
- Elliott, K. C., & Willmes, D. (2013). Cognitive attitudes and values in science. *Philosophy of Science*, 80(5), 807–817.
- Flato, G. M. (2011). Earth system models: An overview. *WIREs Climate Change*, 2(6), 783–800.
- Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy*, 71(1), 5–19.
- Giere, R. N. (2010). An agent-based conception of models and scientific representation. *Synthese*, 172, 269–281.
- Gijsbers, V. (2013). Understanding, explanation, and unification. *Studies in the History of Philosophy of Science Part A*, 44(3), 516–522.
- Greco, J. (2010). *Achieving knowledge. A virtue theoretic account of epistemic normativity*. Cambridge: Cambridge University Press.
- Grimm, S. R. (2010). The goal of explanation. *Studies in the History and Philosophy of Science Part A*, 41(4), 337–344.
- Grimm, S. R. (2017). Understanding and transparency. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 212–229). New York: Routledge.
- Hills, A. (2016). Understanding why. *Noûs*, 50(4), 661–688.
- Hirsch Hadorn, G., & Baumberger, C. (2019). What types of values enter simulation validation and what are their roles? In C. Beisbart & N. J. Sam (Eds.), *Computer simulation validation. Fundamental concepts, methodological frameworks, and philosophical perspectives* (pp. 961–979). Cham: Springer.
- IPCC. (2013). *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*, ed. by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, et al. Cambridge: Cambridge University Press.
- Kelp, C. (2015). Understanding phenomena. *Synthese*, 192(12), 3799–3816.

- Kelp, C. (2017). Towards a knowledge-based account of understanding. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 251–271). New York: Routledge.
- Khalifa, K. (2012). Inaugurating understanding or repackaging explanation? *Philosophy of Science*, 79(1), 15–37.
- Khalifa, K. (2013a). The role of explanation in understanding. *British Journal for the Philosophy of Science*, 64(1), 161–187.
- Khalifa, K. (2013b). Understanding, grasping and luck. *Episteme*, 10(1), 1–17.
- Khalifa, K. (2017). Must understanding be coherent? In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 139–339). New York: Routledge.
- Kuhn, T. S. (1977). Objectivity, value judgment and theory choice. In *The essential tension* (pp. 320–339). Chicago: University of Chicago Press.
- Kvanvig, J. (2003). *The value of knowledge and the pursuit of understanding*. New York: Cambridge University Press.
- Kvanvig, J. (2009). The value of understanding. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Epistemic value* (pp. 95–111). Oxford: Oxford University Press.
- Laudan, L. (2004). The epistemic, the cognitive, and the social. In P. Machamer & G. Wolters (Eds.), *Science, values, and objectivity* (pp. 14–23). Pittsburgh: University of Pittsburgh Press.
- Lipton, P. (2009). Understanding without explanation. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 43–63). Pittsburgh: University of Pittsburgh Press.
- Newman, M. (2012). An inferential model of scientific understanding. *International Studies in the Philosophy of Science*, 26(1), 1–26.
- Newman, M. (2017). Theoretical understanding in science. *British Journal for the Philosophy of Science*, 68(2), 571–595.
- Potochnik, A. (2015). The diverse aims of science. *Studies in the History and Philosophy of Science Part A*, 53, 71–80.
- Pritchard, D. (2010). Knowledge and understanding. In D. Pritchard, A. Millar, & A. Haddock (Eds.), *The nature and value of knowledge. Three investigations* (pp. 1–88). Oxford: Oxford University Press.
- Strevens, M. (2013). No understanding without explanation. *Studies in the History and Philosophy of Science Part A*, 44(3), 510–515.
- Stuart, M. (2016). Taming theory with thought experiments: Understanding and scientific progress. *Studies in History and Philosophy of Science Part A*, 58, 24–33.
- Suárez, M. (2003). Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science*, 17(3), 225–244.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327–352.
- van Camp, W. (2014). Explaining understanding (or understanding explanation). *European Journal for Philosophy of Science*, 4(1), 95–114.
- van Fraassen, B. C. (1980). *The scientific image*. Oxford: Oxford University Press.
- Weisberg, M. (2013). *Simulation and similarity. Using models to understand the world*. Oxford: Oxford University Press.
- Wilkenfeld, D. A. (2014). Functional explaining: A new approach to the philosophy of explanation. *Synthese*, 191(14), 3367–3391.
- Wilkenfeld, D. A. (2015). MUDy understanding. *Synthese*, 194(4), 1273–1293. <https://doi.org/10.1007/s11229-015-0992-x>.
- Wilkenfeld, D. A. (2017). Understanding without believing. In S. Grimm, C. Baumberger, & S. Ammon (Eds.), *Explaining understanding: New perspectives from epistemology and philosophy of science* (pp. 318–333). New York: Routledge.
- Ylikoski, P. (2014). Agent-based simulation and sociological understanding. *Perspectives on Science*, 22(3), 318–335.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.