

Dimensions of Objectual Understanding

Christoph Baumberger, ETH Zurich, *Christoph.Baumberger@env.ethz.ch*

Georg Brun, University of Berne, *Georg.Brun@philo.unibe.ch*

Abstract In science and philosophy, a relatively demanding notion of understanding is of central interest: an epistemic subject understands a subject matter by means of a theory. This notion can be explicated in a way which resembles JTB analyses of knowledge. The explication requires that the theory answers to the facts, that the subject grasps the theory, is committed to the theory and justified in the theory. In this paper, we focus on the justification condition and argue that it can be analysed with reference to the idea of a reflective equilibrium.

1. Introduction

A relatively demanding form of understanding is of central interest in science and philosophy. We strive to understand climate change through climate models, issues in medical ethics by means of a system of ethical principles, or a financial crisis with the help of economic theories. In all these examples, we seek not only a single explanation of a phenomenon, but strive to understand a subject matter (or “domain” or “topic”) by means of a theory. This we call “objectual understanding”.¹

Recent debates about the nature of understanding have mostly been focused on explanatory rather than objectual understanding.² As a consequence, not only have the conditions for ascribing objectual understanding not been analyzed sufficiently, but the very notion of objectual understanding has remained somewhat unclear. It is therefore not surprising that some authors are skeptical whether there is a genuine phenomenon of objectual understanding as opposed to other forms of understanding and knowledge (cf. Khalifa 2013). Only two influential accounts have been proposed up to now. Since Kvanvig (2003; 2009) does not discuss the role of theories as means of understanding, we rather take Elgin’s (e.g. 1996; 2006) account as a starting point and assume that objectual understanding is a distinctive form of understanding.

The goal of this paper is to systematically explore the various dimensions of objectual understanding. In particular, we develop an explication³ of the notion of objectual understanding. This calls for stating conditions which capture key features of our ordinary use of “understanding” in relevant contexts. And the resulting conception should be suitable for some of the roles philosophers have assigned to the notion of understanding. Such roles include making sense of science in general (Elgin 2006; de Regt et al. 2009) and of explanation in particular (Friedman 1974; Grimm 2010), avoiding the value problem for knowledge (Kvanvig 2003; Pritchard 2010; Grimm 2012), identifying intellectual virtues (Zagzebski 2001; Riggs 2003) and characterizing morally worthy action (Hills 2010). Discussing such applications of our proposal, however, lies outside the scope of this paper.

The explication we propose (in Sect. 2) has certain similarities to well-known justified-true-belief accounts of knowledge. It specifies four conditions for a subject to have understanding of a subject matter by means of a theory: that the theory answers to the facts, that the subject is justified in the theory, that it is committed to the theory and that it grasps the theory. Since the first and the last dimension – the theory’s relation to the facts and the subject’s grasping of the theory – have been widely discussed (especially for explanatory understanding), we focus on the remaining two aspects and specifically address the question of what justification amounts to in case of objectual understanding. Our analysis simultaneously helps to clarify the notion of commitment that figures in the explications we suggest.

Our main thesis with respect to the justification condition is that it can be spelled out with the help of the idea of a reflective equilibrium. This proposal is inspired by Elgin’s idea of using a conception of reflective equilibrium as a general account of epistemic justification. The thesis is motivated in Section 3 and systematically explored in Section 4, in which we develop an account of the various aspects involved in the justification of theories. Section 5 considers a range of objections, and Section 6 links the proposed account of justification to the explication of objectual understanding suggested in Section 2. We show how the various conditions for justification relate to the elements of our explication. Together, they characterize the various dimensions of objectual understanding.

2. An explication of objectual understanding

2.1. The target concept

A range of preliminary clarifications are necessary before we can suggest an explication of objectual understanding. We have already restricted the use of “objectual understanding” to an epistemic subject’s understanding of a subject matter by means of a theory.⁴ This sort of understanding firstly contrasts with what is sometimes called “explanatory” or “atomistic” understanding, namely understanding why something is the case by means of an explanation, for example, understanding why the global mean temperature has increased rapidly in the last decades. Secondly, in the literature, “objectual understanding” is also used in a wider, grammatically motivated, sense for other forms of understanding, either for understanding of other things than subject matters (e.g. understanding a specific action of somebody), or for understanding by other means than theories (e.g. by diagrams). Finally, understanding by means of a theory has to be distinguished from understanding a theory (e.g. understanding combustion vs. understanding oxidation theory).

Both the object of objectual understanding – a subject matter – and the means by which it is achieved – a theory – need further clarification. Concerning the object, one might argue that objectual and explanatory understanding differ only in degree because the object of explanatory understanding is a state of affairs (e.g. that the global mean temperature has rapidly increased), and a subject matter (e.g. climate change) is simply a system of states of affairs or a very complex state of affairs. On this basis, Grimm (2011; cf. 2016) argues that the difference between objectual and explanatory understanding is spurious. But this conclusion seems too hasty because objectual and explanatory understanding are also distinguished in terms of the means by which they are achieved. Now theories enable explanations, but they are not merely sets or systems of explanations. Hence, even if subject

matters are simply complex states of affairs, this does not imply that the distinction between objectual and explanatory understanding is spurious.

The notion of a theory used in our characterization of objectual understanding has to be taken in a very wide sense, as including scientific theories in a narrow sense, theory-like representations such as mathematical models, as well as non-empirical theories, normative theories and philosophical theories. Furthermore, we do not presuppose a specific account of the nature of theories. For reasons of simplicity, we talk as if theories were systems of propositions and ignore non-propositional elements, such as categories, diagrams and graphs.⁵

2.2. Necessary conditions for understanding

Two examples will help to show what is involved in objectual understanding.

Example 1. Scientist *S* claims to understand climate change by means of a climate model which includes equations derived from physics with the help of empirical assumptions about cloud albedo and other factors. Defending this claim involves showing several things, both with respect to the model and with respect to the epistemic subject. The model needs to be adequately tied to the facts, and *S* must accept the model since it would be odd to say that she understands climate change by means of a model she is not at least committed to or she even rejects. Furthermore, *S* does not only need to know about the model, but she must have some familiarity with handling the model; more specifically, she needs to grasp the model in the sense of being able to use it in, for example, pre- and retrodictions, scenario analysis, projections (i.e. predictions relative to scenarios) and explanations. Finally, defending *S*'s claim also requires that reasons can be provided which sufficiently speak in favor of the model and show that it is sufficiently good with respect to accommodating the evidence, precision of results, explanatory power and so on – in contrast to flawed accounts which, for example, appeal to changes in solar radiation instead of greenhouse gas emissions.

Example 2. Philosopher *P* claims that she understands issues of medical ethics (e.g. questions about euthanasia, participation in double-blind studies and allocation of donated organs) by means of an ethical theory which includes, say, principles of fair distribution, autonomy, beneficence and nonmaleficence. Again, defending this claim does not only require to show that these principles are right, but also that *P* is committed to them and grasps the relevant theory, which in this case means that *P* is able to provide moral judgments and explanations in a variety of actual and counterfactual cases, for example, to explain why certain distributions of health resources are morally problematic. And again, defending *P*'s claim requires that reasons can be provided which sufficiently speak in favor of the theory and show why it is sufficiently good with respect to intuitive plausibility, providing effective action guidance, explanations why certain actions are objectionable, and so on; in many contexts, we additionally require that the theory in question is at least as good as its competitors which appeal, for example, to some utilitarian calculus, or to considerations of moral virtue.

These examples motivate an explication with four necessary conditions. A grasping condition is needed since understanding by means of a theory requires understanding the theory sufficiently, so

that one is able to make use of the theory.⁶ A justification condition is needed since *S* cannot understand a subject matter by means of a theory if there are not reasons which (with respect to *S*) sufficiently speak in favor of the theory. A commitment condition and a rightness condition are needed since one cannot understand a subject matter by means of a theory one does not accept, nor by a theory which does not answer to the facts or get things right in some other way. As a result, we get the following explication:⁷

A subject *S* understands a subject matter *M* by means of a theory *T* to the extent that

- | | |
|--|----------------------------------|
| (1) <i>S</i> grasps <i>T</i> ; that is, <i>S</i> is able to make use of <i>T</i> | <i>(grasping condition)</i> |
| (2) <i>S</i> is committed to <i>T</i> | <i>(commitment condition)</i> |
| (3) <i>T</i> answers to the facts (or gets things right in some other way) | <i>(rightness condition)</i> |
| (4) <i>S</i> is justified in <i>T</i> of <i>M</i> | <i>(justification condition)</i> |

This explication has some structural similarity to JTB accounts of knowledge. (2) corresponds to the belief condition, (3) to the truth condition and (4) to the justification condition for knowledge. The grasping condition (1) is specific for understanding. We distinguish between the commitment condition and the grasping condition because one can grasp a theory, say, phlogiston theory, which one knows to be inaccurate. Conditions (2)–(3) differ in crucial respects from the corresponding conditions for knowledge. The rightness condition does not require that the theory is strictly speaking true, but allows some peripheral falsehoods (Kvanvig’s quasi-factivity; 2003:201–2) or even some central falsehoods as long as the theory as a whole answers to the facts (Elgin’s non-factivity; 2009:324–7). “Getting things right in some other way” is meant to leave room for the idea that normative theories can be objective even if there are no normative facts they can answer to. And since *S* can understand a subject matter by means of a highly idealized theory *S* knows to be false, the commitment condition does not require belief in the sense of taking to be true. In the next section, we show that also the justification condition needs to be construed in a way that differs from the standard epistemological conceptions of justification, and that it cannot simply be adapted from approaches to the epistemic evaluation of theories familiar from philosophy of science. Instead, we propose to rely on the idea of a reflective equilibrium for spelling out the justification condition for objectual understanding.

3. Justification: the challenge

According to almost all conceptions of epistemic justification developed in epistemology, the core idea of justification is that it is truth conducive: justifying a belief means to make a case for it being true or not being false. These conceptions of justification address knowledge and belief, and are therefore not directly applicable to understanding and theories.

Philosophers of science, on the other hand, have extensively debated the epistemic evaluation of theories. Specifically, three theses have long been discussed, albeit not in the context of understanding (see Douglas 2013 for a short overview):

- (i) An epistemic evaluation of theories needs to appeal to a plurality of epistemic goals, such as Kuhn’s (1977:322) accuracy, consistency, broad scope, simplicity and fruitfulness.

(ii) These goals are goals in themselves and not only insofar they are truth conducive (if they are truth conducive at all).

(iii) Trade-offs between these goals are often unavoidable, and in general, there is not a unique best theory.

(i) and (ii) constitute epistemic value pluralism. (iii) is usually part of the package, but it is independent of (i) and (ii) since the different goals could be ordered lexically. All three points have been controversial, but it can be argued that they must be accepted if epistemic evaluation is understood as evaluation with respect to a theory's contribution to understanding. Here are three examples.

Broad scope. Other things being equal, broadening the scope of a theory advances the understanding of the original range of phenomena by showing that they are instances of a more general pattern. In other cases, the advancement of the understanding is due to the ability to show that one of two theories which are formerly thought to be unconnected can be derived from the other. Newton's gravitational theory advanced the understanding of the free fall of a body near the surface of the Earth by explaining it as an instance of a more general pattern which includes such diverse phenomena as the generation of the tides and the movement of celestial bodies (see Hoyningen-Huene 2013:60–1). This example supports (i) for the epistemic evaluation of theories with respect to their contribution to understanding.

Simplicity. Other things being equal, simplifying a theory can enhance understanding even if the simpler theory is equivalent to the more complex one. In propositional logic, the reduction of the logical operators to the Sheffer stroke results in a simpler theory in the sense of a theory with fewer basic notions. Having such a theory enhances our understanding of logical relations even if this theory may be harder to grasp. Since both theories are equivalent, being truth conducive is at least not the whole point of simplicity. Hence, this example supports not only (i) but also (ii) for the epistemic evaluation of theories with respect to their contribution to understanding.

Scope vs. fruitfulness. Understanding may be enhanced by reducing scope if this leads to a more fruitful theory. Excluding the whales from the fish leads to a biological theory of fish with reduced scope but more and simpler regularities and thereby enhances understanding of the animal kingdom.⁸ This example supports (iii); understanding can sometimes be advanced only if a trade-off between epistemic goals is accepted.

These examples show that in the context of objectual understanding, there are theories which must be justified with respect to a plurality of epistemic goals that are not exclusively truth conducive and admit of trade-offs. As a result, standard epistemological accounts of justification cannot be adapted to our project since they conceive of justification as exclusively truth conducive.⁹

However, can we not simply use theory-choice approaches or confirmation theory as accounts of the epistemic justification of theories? After all, we have argued for the unsuitability of standard epistemological accounts of justification by making use of insights from the debate in philosophy of science about the epistemic evaluation of theories. However, solving the problem of justification in this way does not fit our project, given that not only scientific theories, but also non-empirical, normative and philosophical theories can provide objectual understanding. As we will see in Section

4.4, theories of confirmation and the discussion about theory-choice remain relevant in our account, but our account cannot be reduced to them.

One reason why reflective equilibrium is a promising candidate for a general conception of justification for theories enabling objectual understanding is that it is tailored to a plurality of epistemic values which are not exclusively truth conducive and admit of trade-offs. This thesis is only plausible if reflective equilibrium is not identified with coherence.¹⁰ We rather defend a more ambitious account of reflective equilibrium, based on the work of Goodman and especially on Elgin's idea of using a conception of reflective equilibrium as a general account of epistemic justification.

4. A reflective-equilibrium account of justification

Reflective equilibrium was first described in Goodman's (1983) discussion of the justification of logics, but it is best known for its application to moral theories by Rawls (1999a; 1999b) and Daniels (1996). Elgin's project (1996) of a reflective-equilibrium-based epistemology has been the most elaborated account so far. It is also the basis for the account we propose in the following subsections (elaborating on the sketch in Brun 2014).

4.1. Some basic ideas of reflective equilibrium

Two key ideas characterize the method of (so-called "wide") reflective equilibrium: (1) Judgments and principles are justified if judgments, principles and background assumptions are in equilibrium; (2) this state is reached through a process that starts from judgments and background assumptions, proposes systematic principles and then mutually adjusts judgments, principles and possibly also background assumptions.

All standard accounts of reflective equilibrium characterize the contrast between judgments and principles in terms of particular vs. general (e.g. Goodman 1983:64). However, Rawls (1999b:289) pointed out that judgments can also be general, and Goodman in fact contrasts an account of validity given in a logical system, such as *Principia Mathematica*, and the judgments of validity somebody actually forms or is ready to accept. Similarly, the distinction between moral principles and moral judgments can be interpreted as the distinction between what is part of some moral theory and what somebody is actually committed to to a certain degree. In the context of empirical theories, we can distinguish claims that are part of, say, Newton's mechanics or the phlogiston theory of combustion from general or particular propositions referring to mechanical or chemical states of affairs we actually endorse, for example, because they are confirmed by evidence or seem to be true intuitively. In all these cases, the contrast between judgments and principles is not a matter of their content; this can be the same. The crucial difference is rather that principles are part of some theoretical system (of, e.g. logic, morality, mechanics or chemistry), whereas judgments include a propositional attitude that involves some (at least minimal) degree of commitment. To avoid the association with particular vs. general and the identification of judgments with explicit statements, we replace "judgment" and "principle" with the technical terms "commitment" (following Elgin 1996:102–9) and "element of a theory".

Within the process of mutually adjusting theory and commitments, commitments are involved in two contrasts. On the one hand, commitments contrast with the elements of the theory at every stage in the process of developing a reflective equilibrium. On the other hand, the resulting position, i.e. the resulting commitments and theory, contrast with the commitments the reflective-equilibrium process started out with. For the sake of clarity, we refer to the latter as “antecedent commitments” and use “current commitment” or simply “commitment” in the context of the first contrast.

With these basic points in mind, we can now unpack the idea of a reflective equilibrium.

4.2. Conditions for justification

The central metaphor of an equilibrium is usually explained in terms of coherence. And Rawls’s widely quoted remark “[...] justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view” (1999a:19) may have encouraged the view that reflective equilibrium is just a coherentist account of justification (e.g. de Sousa 2010:23). However, coherence of judgments and principles is not all that matters for justification and “reflective equilibrium” is not just another label for coherentism. This becomes clear in a closer analysis which does not confine itself to the routinely quoted passages from Goodman and Rawls.

In what follows, we discuss criteria for justification that can be reconstructed from Goodman’s and Elgin’s writings. Roughly, the idea is this: to be justified in a theory, an epistemic subject must hold a position which includes this theory and which is in reflective equilibrium, and the subject must be able to make sufficiently plausible that this is indeed the case. Being in reflective equilibrium, in turn, is matter of three criteria: that the subject’s position is in equilibrium, that it does justice to epistemic goals and that it respects the subject’s antecedent commitments. Our official formulation includes the complications necessary to make explicit that being in reflective equilibrium must be relativized to a range of factors:

- (I) An epistemic subject *S* is justified in theory *T* of subject matter *M* iff there is a position *P* consisting of a set of commitments *C* and a theory *T* such that *C* is a subset of *S*’s commitments, *P* is in reflective equilibrium (relative to a configuration of epistemic goals *G*, background theories *B*, background information *I*, and *S*’s antecedent commitments *A* about *M*), and *S* can make it sufficiently plausible that *P* is in reflective equilibrium.
- (II) A position *P*, consisting of a set of commitments *C* and a theory *T*, is in reflective equilibrium, relative to a configuration of epistemic goals *G*, background theories *B*, background information *I* and antecedent commitments *A* iff
 - (i) *C*, *T*, *B* and *I* are in equilibrium
 - (ii) *P* does justice to *G*
 - (iii) *C* respects *A*

We give two clauses because this enables us to disentangle those aspects of justification which are essentially tied to an epistemic subject from those which are not. According to (II), being in reflective equilibrium is an attribute of a position; (I) then makes the link to the epistemic subject. Before we discuss conditions (i)–(iii) in the following subsections, some comments on (I) and (II) are needed.

Firstly, to keep things reasonably simple, (I) and (II) treat justification and being in reflective equilibrium as yes-no matters. However, it obviously makes sense to speak of being more or less justified, of being more or less in reflective equilibrium, and of meeting conditions (i)–(iii) to some greater or lesser extent. Although we will not go into the details of how to adapt our proposal by introducing comparative concepts of justification etc., we must briefly discuss how we should deal with the comparative character of justification in the context of an explication of objectual understanding. Specifically, should we introduce a non-comparative notion of being justified by introducing one or both of the following necessary conditions: a position is justified only if it is in reflective equilibrium *to a sufficient degree*; a position is justified only if there are no relevant alternative positions which are more in reflective equilibrium (with respect to the same epistemic goals *G*)?

Both conditions are plausible in the context of the question “under what conditions is it all things considered reasonable to adopt a position?” It then clearly does not suffice to answer that the position has just some, possibly minimal, degree of justification, and it is implausible that a position should be adopted even if better justified alternatives are available.

The context of objectual understanding, however, is less demanding. It would still be odd to say that one can understand a subject matter by a theory which is only minimally justified. Clearly, we must require that the theory enjoys a sufficient degree of justification, although it is less clear how we could spell out what exactly is required for a sufficient degree. However, text-book theories and examples from the history of science show that objectual understanding does not always need a theory which is at least as well justified as available alternatives. We can still gain some understanding of chemistry by means of old and simple models appealing to valence bonds or electron shells.¹¹ Of course, in other contexts we may set higher standards and want to understand something by means of the best theory available (or at least a maximally good theory; cf. the examples in Sect. 2.2).

Secondly, (I) explicitly restricts the antecedent commitments to those of *S*, but it leaves it open whose epistemic goals *G*, background theories *B* and background information *I* count. In the context of objectual understanding, *G*, *B* and *I* should be understood as related to *S*. A subject can have some understanding of a subject matter if she grasps and accepts a correct theory which is part of a position that is in reflective equilibrium relative to *her* epistemic goals, background theories, background information and antecedent commitments.¹² However, in the context of theory acceptance, we ask under what conditions it is *all things considered* reasonable to adopt a certain position, and therefore we require that this position is in reflective equilibrium relative to the best available background theories and to all available background information that is relevant for the questions at issue.

4.3. Equilibrium

Explaining the metaphor of equilibrium in terms of coherence introduces at least three requirements: the commitments and the theory must be consistent; it must be possible to infer the commitments from the theory, and the background theories must support the theory.¹³

The relation of inferability is normally not a matter of straightforward deductive consequence, but often involves plausible reasoning, and nearly always a transition from a more or less formal theory to ordinary language expressions of commitments. Typical examples for the first point are moral theories

that require reasoning about which principles are most relevant in the case at hand or which reasons prevail all things considered (see, e.g. the model of Beauchamp/Childress 2013). The second point is most easily seen in case of logical systems, which are framed in a formal language, not the language we use to express our commitments. Moral theories are typically presented in English, but their relation to commitments is usually not trivial either because moral theories use ordinary words in technical ways.

Moreover, many of the most relevant commitments cannot be inferred from the theory alone, but only with the help of some background information. In the context of moral issues, such information could include how much pain a patient is suffering and whether she is deciding under pressure of her doctors. In case of climate models, two important types of background information are boundary conditions, e.g. assumptions about emission patterns of greenhouse gases, and initial conditions, e.g. empirical data about current temperature.

Background theories, finally, are theories which support the theory in the foreground. Examples are basic theories from physics and chemistry which provide the theoretical background for climate models, and psychological theories of motivation and decision making which are appealed to in moral theorizing.

The distinction between background and foreground is not intended as a distinction between information or theories belonging to different subject matters. It is rather a question of perspective intended to capture the fact that inquiry has to proceed piece-meal rather than wholesale: while we seek to give a justification for the position in the foreground, we appeal to further assumptions, which are in the background insofar as we treat them as independently justified. Of course, background theories and background information need to be justified as well, but their justification is not in focus as long as they are in the background. In the example of climate models, justification of the model is in focus, particularly equations adopted (with some modifications) from physics and chemistry as well as parameters such as empirical assumptions about cloud albedo. What is not in focus is the justification of the chemical and physical background theories, as well as the justification of the boundary and initial conditions which constitute the background information necessary for inferring empirical statements from the model.

In sum, we get the following official condition:

- (i) Commitments C , theory T , background theories B , and background information I are in equilibrium iff
 - (a) C and T are consistent
 - (b) C and only C can be inferred from T with the help of I
 - (c) T is supported by B

To see why the equilibrium condition is necessary for justification in the context of objectual understanding, let us have a closer look at what the above conditions demand. We can isolate three necessary conditions for understanding a subject matter by means of a theory:

- (a) Internal consistency of the theory; that is, one cannot understand by means of an inconsistent theory.

- (b) Internal coherence of the position: the theory fits together with the subject's commitments. Specifically, one cannot understand by means of a theory one does not accept.
- (c) External coherence of the position: the theory and what can be inferred from it is supported by the subject's background theories. One cannot understand by means of a theory which is incompatible with one's background theories. At the very minimum, the theory must be consistent with the relevant background theories. Moreover, external coherence can positively enhance understanding. While it may be possible to gain some understanding by means of an "isolated" theory which is merely consistent with but mostly logically independent of relevant background theories, further support by background theories boosts understanding by integrating it into a wider picture.

4.4. Doing justice to epistemic goals

The requirement that a position in reflective equilibrium must meet a configuration of epistemic goals should secure at least two things: that the position is credible and that it is systematic. "Credibility" is what provides reasons to assume that a theory answers to the facts or is right in some other way; hence, that it meets the rightness condition (from Sect. 2.2). In what follows, we discuss a range of epistemic goals and what they contribute to credibility and systematicity.

If a position is to be credible, at least some of the commitments it includes must have a minimal epistemic standing of "independent credibility"; that is, credibility which is independent of the current equilibrium, in particular, independent of the commitment's coherence with other commitments of the position and with the theory (Goodman 1972:62–3; Rawls 1999b; cf. Elgin 1996:101–7). Without independently credible commitments, the method of reflective equilibrium would be susceptible to the standard objection that coherence cannot generate justification *ex nihilo*.¹⁴

Like credibility in general, independent credibility comes in degrees and may be minimal only. But even a high degree of independent credibility is not sufficient for justification. Since independent credibility is by definition independent of the equilibrium in question, it cannot be sufficient for justification if reflective equilibrium is necessary. Furthermore, commitments with any degree of independent credibility are still commitments in the technical sense and as such not immune from revision.

Three reasons for ascribing independent credibility to a commitment are worth pointing out. Firstly, a commitment may be independently credible because it accommodates some evidence. In case of empirical theories, such credibility may, for example, be attributed to commitments referring in a relatively direct way to data gained by observation, or to commitments based on testimony. For non-empirical theories, many philosophers (paradigmatically Bealer 1996) hold that intuitions play the role of evidence. Intuitive commitments may refer to conceptual relations, to the validity of arguments or to substantial matters such as the moral (im)permissibility of (types of) actions. Investigating which commitments in fact have independent credibility is a task not for the reflective-equilibrium account of justification, but for theories of confirmation, perception and testimony, for example.¹⁵ The reflective-equilibrium account just shows which role such epistemic theories play when it comes to the actual justification of a specific theory.

Secondly, the commitments which we start out with must have some independent credibility. If they had no credibility at all, they would not qualify as commitments. In the limiting case, the basis for

ascribing some independent credibility to an antecedent commitment is just that so far nothing speaks against entertaining this commitment. And since such commitments are antecedent relative to the reflective equilibrium we are seeking to develop, their credibility can only be independent. On this basis, independent credibility can also be ascribed to those commitments which are both antecedent and current commitments in the resulting position.

Finally, background assumptions bestow independent credibility to the commitments they support. In the context of justifying a theory of distributive justice, a commitment to help people which are much less well-off may be supported by a moral background theory which includes a requirement to help the poor.

Further epistemic desiderata can be divided into generally relevant virtues and virtues which are specifically relevant to certain kinds of theories. Generally relevant are, for example, precision, simplicity, conceptual clarity, fruitfulness, explanatory power, completeness with respect to a subject matter, and broad scope of application. In the literature on theory-choice, consistency and empirical accuracy are often listed as well (e.g. Kuhn 1977:322). In our account, consistency is given a special place – it is mentioned not in condition (ii), but in condition (i) – because it is a key ingredient of equilibrium and does not admit of trade-offs.¹⁶ And empirical accuracy is treated as a specific variant of the more general requirement of credibility, which also covers non-empirical theories. In any case, credibility does not reduce to empirical accuracy because in addition to accommodating evidence further virtues of theories can provide reason to assume that it is true, for example, explanatory power.

One may wonder why doing justice to such epistemic desiderata should count as an aspect of reflective equilibrium at all. A first answer is that at least some desiderata contribute to a theory's credibility. Secondly, one of the basic ideas of the method of reflective equilibrium is that there should be an agreement between commitments and a theory; that is, a *systematic* set of principles. Doing justice *simultaneously* to desiderata such as precision, simplicity and wide scope ensures that the set of principles in fact has some degree of systematicity. Finally, the desiderata also block the “conservative” strategy of minimizing revisions by selling a streamlined list of antecedent commitments as a “theory”.

In addition to generally relevant desiderata, there are epistemic desiderata which are specifically relevant to certain kinds of theories and their justification. In case of scientific theories, candidates for more specific virtues include visualizability and causality (cf. de Regt/Dieks 2005); for moral theories, being an effective guide to action or being a suitable basis for legal regulation may count as important virtues.

Independent credibility and the various desiderata are bundled in one criterion because they are all epistemic goals and because doing justice to a plurality of epistemic goals can involve trade-offs between any of them. Increasing the simplicity of a theory may only be feasible by discarding some commitments with independent credibility and thereby rejecting pieces of evidence. On the other hand, maintaining credibility blocks oversimplifications and sweeping generalizations one may be tempted to accept in the name of systematicity. That trade-offs are typically unavoidable is also a reason why we speak of “doing justice” rather than “realizing” epistemic goals. Although justification calls for taking epistemic goals seriously it would be unrealistic to insist on theories which effectively reach all those goals simultaneously.

Which desiderata are relevant, how much weight they should be given, and which trade-offs are acceptable depends on the subject matter and on the specific pragmatic-epistemic goals which guide the construction of the theory at hand. This is the reason why we speak of a *configuration* of epistemic goals. If, for example, a climate model is developed for the purpose of understanding the basic mechanisms (pragmatic-epistemic goal) of global climate change (subject matter), the model should be as simple as possible and it may be a decisive advantage if it can be visualized, whereas we may not insist on it being useful in effectively calculating exact figures for the key factor it involves. If, on the other hand, we want to understand regional climate change in a way that makes reliable predictions or projections of temperature, precipitation, etc. available, we will require that the model be as detailed as necessary for effectively computing the relevant factors with sufficient precision even if this means that the model gets incredibly complicated.

The discussed plurality of epistemic goals also sheds some light on the notion of commitment, which plays a central role in our account of reflective equilibrium. Commitment is an epistemically relevant status which is not just related to credibility and truth (as, e.g. belief), but to a whole plurality of epistemic goals which contribute to a theory's credibility and systematicity.

The official formulation of the condition is:

- (ii) A position P consisting of a set of commitments C and a theory T does justice to a configuration of epistemic goals G iff some commitments in C have independent credibility and T does justice to the further desiderata in G .

This condition fits particularly well into a conception of justification for objectual understanding since it accounts for the points made in Section 3: it refers to a plurality of epistemic goals which allow trade-offs and cannot all be explained as exclusively truth conducive because they can be in competition with credibility. Moreover, (ii) can account for the fact that the standards a justification must meet are not just given. Rather, the pragmatic-epistemic goals of understanding (partly) determine the configuration of epistemic goals to which a theory must do justice.¹⁷ Finally, reflective equilibrium is especially at home in an account of understanding since the method is plausible only if we accept that there is plurality of epistemic goals. Otherwise we would be better off saying that credibility is all that matters for justification. However, this plurality raises the question of which goals should count as epistemic. The question becomes especially pressing if we reject the claim that truth is the primary epistemic goal and all secondary goals are *epistemic* only insofar as they are truth-conducive. A promising answer refers to understanding: epistemic goals are those goals which are relevant for understanding.

However, this explanation of condition (ii) also invites the objection that it is circular to define objectual understanding (indirectly) in terms of epistemic goals and to characterize epistemic goals in terms of understanding. The key to the answer is that we deal with an *explication* of understanding. What is defined in terms of epistemic goals is a theoretical notion of understanding, whereas the notion of understanding which is used to characterize epistemic goals is our colloquial concept of understanding. The former is the explicatum, the latter the explicandum. Hence no circularity results because two concepts of understanding are involved. That the explicandum guides the account we give of the explicatum is just a characteristic trait of explication.

4.5. Respecting antecedent commitments

Reflective equilibrium also demands that the resulting position must respect antecedent commitments adequately. This condition prevents the process of developing an equilibrium from implementing revisions so drastic that it “changes the subject”; that is, that we end up with a system that does not count as a theory of what we set out to develop a theory of. If a course of reflection leads to a theory which merely underwrites the law of the jungle, it will not count as providing a *moral* theory since it would force us to give up too much of our most important moral commitments. Similarly, if a model turns out to describe only short-term conditions of meteorological variables such as temperature and precipitation in a given region, it will not count as a *climate* (in contrast to *weather*) model.

Sources of antecedent commitments range from merely plausible assumptions to commitments that are based on a previous theory which is now to be revised. In any case, antecedent commitments must have some independent credibility: *qua* commitments, they must have some credibility, and *qua* antecedent, their credibility can only be independent.

Respecting antecedent commitments means that we must be ready to explain why we have discarded, replaced or modified certain antecedent commitments. Typical explanations refer to the relative weight of commitments and elements of the theory, or they point out a difference between two commitments which motivates us to save the one but give up the other, or they provide a diagnostic story which explains why we were committed to something we no longer are. Here is an example from moral philosophy: If a supporter of gay marriage who appeals to a “between consenting adults” principle finds that he cannot uphold his commitment that polygamy is morally wrong, he can argue that the latter was feeble anyway in comparison to the principle which covers many commitments he firmly holds or that his views on polygamy were merely a cultural prejudice.

Here is the official version of the condition:

(iii) The commitments *C* respect the antecedent commitments *A* iff a reasonable explanation can be given why those commitments which are in *A* but not in *C* have been given up or altered.

This criterion gets additional motivation in the context of objectual understanding: gaining a better understanding of a subject matter often involves adopting a new theory, which, in turn, often requires considerable re-thinking. In such cases, it is an open question whether one indeed understands what one tried to understand. In case of empirical theories, the question is especially pressing in connection with paradigm changes. Similarly for philosophical theories. Does, for example, a transition from an utilitarian to a Kantian ethics change the subject matter? Furthermore, many philosophers seem to be committed to the position that only very restricted alterations of commitments are admissible because they hold that theories must virtually always be revised if they face counter-examples. One reason for such a stance may be the fear of inadvertently changing the subject matter. On the other hand, philosophers who reject a counter-example in favor of a theory must also be ready to answer the challenge that their theory in fact deals with another subject. In all these cases, some argument why theory *T* does not change the subject matter *M* is relevant for the justification of *T as a theory about M*.

4.6. The process of adjusting commitments and theory

As a second key idea, the method of reflective equilibrium includes a process of mutual adjustment of commitments and theories. Speaking of *mutual* adjustments does not only emphasize bidirectionality, but also underlines that neither the elements of the theory nor the commitments have *as such* a privilege not to be revised. A second important point is that, as Goodman (1983:65) insisted, the process of mutual adjustment is neither intended to describe how we in fact arrive at a position, nor is it to be read as a recipe one should follow in developing a position. It rather spells out what is required for a justification: we need to be able to describe how the theory could have been developed from the commitments we started out with, irrespective of whether we have actually devised the theory in this manner. Hence, the process is a reconstruction (in the sense of Carnap 1963:16).¹⁸

However, if the process is a reconstruction, one may wonder what this procedural aspect contributes to the resulting conception of justification. Has adjusting commitments and theory independent justificatory relevance? Should it be a necessary condition for a position to be justified that it can be reconstructed as the result of such a process? It is hard to see what running through a process of mutual adjustments in itself contributes to justification. It is more plausible to argue two other points, which make the process of adjustments *indirectly* relevant to justification. Firstly, finding a position in reflective equilibrium may often be possible only through a step-by-step process in which a theory is developed and successively adjusted with the relevant commitments. Secondly, given a position, a configuration of epistemic goals and a set of antecedent commitments, there may be no way to directly tell whether the criteria for a reflective equilibrium are met. Specifically, it may be impossible to say whether a reasonable trade-off between the various epistemic goals has been made, or whether a position which respects more (or less) antecedent commitments but does less (or more) justice to the epistemic goals would meet the criteria for being in reflective equilibrium better all things considered. It may well be that the only way to answer such questions is to argue that the position in question could have been reached by a process consisting of adjustments which can all be defended as reasonable steps towards a state of reflective equilibrium.

4.7. The internalist requirement

As an account of justification that is plausible for objectual understanding, the conditions discussed so far need to be supplemented by an internalist requirement.¹⁹ This can be shown by some kind of guessing cases. Suppose *S* borrows a book about a subject matter in which she is a complete novice, say, counter-point; *S* does not even consider whether the author is reliable, but nonetheless accepts the theory presented in the book; and in fact, the acquired position is in reflective equilibrium, relative to *S*'s (minimal) antecedent commitments about counter-point, the relevant configuration of epistemic goals, background theories and background information. However, if we can say no more about *S* than that she has accepted such a theory, this seems clearly an insufficient basis for claiming that *S* is justified in *T*. What *S* does just boils down to a complicated and time-consuming version of guessing.

In order to deal with such cases, (I) includes the internalist requirement that *S* can make it sufficiently plausible that her position is in reflective equilibrium.²⁰ This does not demand that *S* be able to show conclusively that her position is in reflective equilibrium; for example, she does not need to be able to actually infer her commitments by way of explicit arguments from her theory. She only

needs to be able to make it sufficiently plausible that her theory and her commitments are in reflective equilibrium. Furthermore, the internalist requirement is open to a contextualist interpretation; in a scientific context, for example, more is required of an epistemic subject than in an everyday context.

5. Objections

5.1. Are the conditions for justification too demanding?

There is the general worry that the justification conditions as specified in Section 4 are too demanding since practical chances to reach a reflective equilibrium are often dim (cf. DePaul 2011:cii). To assess this worry, it is important to be clear how to understand it. Here are three possibilities; the third is the most serious one.

Firstly, the worry could be that it is virtually impossible to develop any position in reflective equilibrium. But there clearly are positions, such as classical propositional and first-order logic, which are (sufficiently) in reflective equilibrium.

Secondly, the worry might be that at least in some domains it is utterly unclear whether and how one could reach a reflective equilibrium since all available theories either have consequences we are not willing to accept, perform badly with respect to epistemic goals, or require us to give up too many of our antecedent commitments. Ethics might provide an example. Utilitarian theories perform well with respect to some epistemic goals (e.g. they are quite simple and precise and have broad scope), but they have consequences we may not be willing to accept (e.g. torturing innocent people is permitted if the life of a greater number can be saved thereby). Other theories might avoid some of these consequences, but only at the cost of other counterintuitive results (Kantian ethics, e.g. disallows lying even in cases where most of us would recommend it) and sometimes also at the cost of doing badly with respect to epistemic goals. However, even if there are domains in which no position is in reflective equilibrium,²¹ this does not show that our account of justification is too demanding. It only shows that there are domains in which it is difficult to develop justified theories and which are therefore difficult to understand.

Thirdly, the worry could be that an epistemic subject *S* is justified in theory *T* even though *S*'s position is not in reflective equilibrium. Most promising candidates for such a situation are cases in which condition (i.b) – the commitments can be inferred from the theory – is not fulfilled,²² either (1) because *S* does not commit herself to everything that can be inferred from *T*, or (2) because some of *S*'s commitments are incompatible with what can be inferred from *T*. However, it is quite implausible that such cases are counterexamples to our account. It is clearly odd to say that a subject *S* is justified (in the non-technical sense of this word) in a theory *T* or understands a subject matter by means of *T* even though *S* does not accept *T*, or even more implausibly, even though *S* has commitments that are incompatible with *T*.

A more detailed answer with respect to type-(2) cases elaborates on the conditions for a commitment to be incompatible with something that can be inferred from *T*. In short: if not every inconsistency whatsoever in *S*'s commitments about any subject matter should have the result that all of *S*'s commitments are incompatible with some other commitment of *S* (because of the explosion property of logical consequence), we must restrict the incompatibilities relevant to type-(2) cases.

Relevant cases are those in which *S* acknowledges a commitment which is directly (independently of any other commitments) incompatible (at least logically contrary) with something that can be inferred from *T*. In such cases, it is even more implausible as in type-(1) cases that *S* is justified in *T*.

5.2. Are the conditions for justification too weak?

On the other hand, there is also the worry that the proposed justification conditions are not demanding enough since there seem to be cases in which all conditions are fulfilled but we are not willing to accept that the position is justified for the epistemic subject. Specifically, it may seem that the method of reflective equilibrium can be used to “justify” clearly false positions if an epistemic subject starts out with antecedent commitments and background theories that are simply false. Call this the “garbage in, garbage out” objection. A standard case is an epistemic subject who stubbornly appeals to the gambler’s fallacy. The gambler has a strong commitment to a claim such as:

(G1) Black is more likely to show up after red has prevailed for some time.

The gambler infers this commitment from some faulty principles of probability such as (G2) and the background information that (in a fair gamble of Roulette) black and red are equally probable.

(G2) If two types of events are equally likely to occur, then events of type 1 are more likely after a series of events of type 2.

Now the gambler may be in reflective equilibrium, even though (G1) and (G2) are simply false; but it certainly is implausible to say that the gambler is justified in his theory (Stich/Nisbett 1980). This objection can be answered in two closely related ways. Firstly, the case loses some plausibility if we insist on including background theories in the description of the case. If the gambler accepts standard probability theory, he faces stock arguments which show that this background theory is incompatible with (G2); hence we do not have a state of reflective equilibrium. Consequently, a convincing case must not include any background theories incompatible with (G2). This means that subjects with sufficiently weird antecedent commitments must subscribe to equally weird background theories or avoid background theories which are incompatible with their commitments.²³ However, if the gambler does not accept standard probability theory, appealing to a social aspect of inquiry provides a second answer. The gamblers of our example know perfectly well that other epistemic subjects disagree with their commitments to (G1) and (G2) as well as with their background theories. But having such testimony constitutes further evidence the gamblers must accommodate if their positions are to be in reflective equilibrium. The result is even more pressure to revise commitments or principles such (G1) and (G2).²⁴

In the context of objectual understanding, there is still another way of dealing with epistemic subjects who are ready to pay any epistemic price whatsoever as long as they can cling to their weird commitments. Irrespective of whether such subjects are justified in their weird theories, they cannot understand any subject matter by means of these theories since they do not meet the rightness condition.

6. Conditions for objectual understanding

How does the reflective equilibrium account of justification fit into our explication of objectual understanding? More specifically, how is it related to the grasping condition and the commitment condition given in Section 2.2? The relation to the grasping condition depends on how the latter is spelled out. Following de Regt (2009), we suggest that grasping a theory can be explained as being able to make use of it. According to many theorists (e.g. Hills 2010:ch. 9, de Regt 2009), this ability, in turn, involves the ability to apply the theory to actual and counterfactual cases by inferring conclusions from the theory and giving explanations with its help.²⁵ It is debated whether the ability to give explanations is really necessary (e.g. Lipton 2009). However, even a grasping condition which requires the ability to explain does not imply the justification condition, since a subject can grasp a theory she does not accept so that the theory and her commitments are not in equilibrium. On the other hand, the justification condition requires the ability to make it sufficiently plausible that a position is in reflective equilibrium and hence the ability to draw conclusions about actual and counterfactual cases from the theory. This raises the question of whether these abilities in turn imply the ability to use the theory in explanations. The answer depends on what exactly is needed for giving explanations. If not every correct prediction provides an adequate explanation, one can have these abilities without being able to use the theory in explanations. To leave room for such a view, the explication of objectual understanding needs a grasping condition in addition to the justification condition.

The commitment condition does not imply the justification condition. A subject can be committed to everything which can be inferred from a theory even if her position is incompatible with background theories, or does not justice to the relevant configuration of epistemic goals, or does not respect her antecedent commitments. This is how it should be, since we clearly can be committed to theories we are not justified in. However, the justification condition implies the commitment condition, since being justified in a theory implies that the theory is part of a position which is in reflective equilibrium and this, in turn, implies that the subject is committed to everything which can be inferred from the theory and, in this sense, committed to the theory. This makes the commitment condition redundant, but it is nonetheless reasonable to stick to it in the explication of objectual understanding because commitment were necessary for understanding even if we would spell out the justification condition in a way which does not imply the commitment condition.

7. Summary

We have explored various dimensions of a relatively demanding notion of objectual understanding: an epistemic subject understands a subject matter by means of a theory. For this notion we have first suggested an explication which resembles JTB analyses of knowledge. The explication requires that the theory answers to the facts, that the subject grasps the theory, is committed to the theory and justified in the theory. We have then focused on the justification condition and argued that it can be analyzed with reference to the idea of a reflective equilibrium.

This approach acknowledges that the epistemic evaluation of theories is not exclusively related to their truth, but relates to a plurality of epistemic goals which can be subject to trade-offs. The basic idea of reflective equilibrium is that having a justified theory amounts to having a credible position

which includes a systematic theory about the subject matter at hand. This idea motivates three conditions: (i) the subject's commitments, the theory and relevant background theories must agree with one another; (ii) the theory must do justice to epistemic goals, which ensure its credibility and systematicity; and (iii) the resulting position must respect the subject's antecedent commitments about the subject matter, which ensures that the theory is in fact a theory about this subject matter. In the context of objectual understanding, justification also requires that the epistemic subject can make it sufficiently plausible that her position is in reflective equilibrium.

Our analysis of the justification dimension of objectual understanding simultaneously helps to clarify the notion of commitment that figures in the suggested explication. Commitment is an epistemically relevant status which comes in degrees, can be minimal or feeble, is not immune from revision and not just related to credibility and truth (as, e.g. belief), but to a plurality of epistemic goals which contribute to a theory's credibility and systematicity.

Acknowledgments

Earlier versions of this paper have been presented in Berne, Karlsruhe and Osnabrück. For helpful discussions and critical comments, we thank the audiences and in particular Claus Beisbart, Gregor Betz, Catherine Elgin, Kareem Khalifa, Jörg Löschke and Sebastian Schmoranzer. Georg Brun's research is funded by the Swiss National Science Foundation (project 150251).

References

- Baumberger, Christoph; Claus Beisbart; Georg Brun. 2016. "What is Understanding? An Overview of Recent Debates in Philosophy". In Grimm, Stephen; Christoph Baumberger; Sabine Ammon (eds). *Explaining Understanding. New Perspectives from epistemology and Philosophy of Science*. New York: Routledge. xx–yy.
- Bealer, George. 1996. "A Priori Knowledge and the Scope of Philosophy". *Philosophical Studies* 81, 121–142.
- Beauchamp, Tom L.; James F. Childress. 2013. *Principles of Biomedical Ethics. 7th ed.* New York/Oxford: Oxford University Press.
- Brun, Georg. 2014. "Reflective Equilibrium without Intuitions?" *Ethical Theory and Moral Practice* 17, 237–252.
- Brun, Georg. 201*. "Explication as a Method of Conceptual Re-Engineering" [<]
- Carnap, Rudolf. 1962. *Logical Foundations of Probability. 2nd ed.* Chicago: University of Chicago Press.
- Carnap, Rudolf. 1963. "Intellectual Autobiography". In Schilpp, Paul Arthur (ed.). 1963. *The Philosophy of Rudolf Carnap*. La Salle: Open Court. 3–84.
- Daniels, Norman. 1996. *Justice and Justification. Reflective Equilibrium in Theory And Practice*. Cambridge: Cambridge University Press.
- DePaul, Michael R. 2011. "Methodological Issues. Reflective Equilibrium". In Miller, Christian (ed.). *The Continuum Companion to Ethics*. London: Continuum. lxxv–cv.

- Douglas, Heather. 2013. "The Value of Cognitive Values". *Philosophy of Science* 80, 796–806.
- Elgin, Catherine Z. 1996. *Considered Judgment*. Princeton: Princeton University Press.
- Elgin, Catherine Z. 2006. "From Knowledge to Understanding". In Herrington, Stephen (ed.). *Epistemology Futures*. Oxford: Clarendon Press, 199–215.
- Elgin, Catherine Z. 2009. "Is Understanding Factive?" In Haddock, Adrian; Millar, Adrian; Pritchard, Duncan (eds). *Epistemic Value*. Oxford: Oxford University Press, 322–330.
- Elgin, Catherine Z. 2014. "Non-foundationalist Epistemology. Holism, Coherence, and Tenability". In Steup, Matthias; John Turri; Ernest Sosa (eds). *Contemporary Debates in Epistemology*. 2nd ed. Malden: Wiley Blackwell. 244–255.
- Elgin, Catherine Z. 2016. "Exemplification in Understanding". In Grimm, Stephen; Christoph Baumberger; Sabine Ammon (eds). *Explaining Understanding. New Perspectives from epistemology and Philosophy of Science*. New York: Routledge. xx–yy.
- Friedman, Michael. 1974. "Explanation and Scientific Understanding". *The Journal of Philosophy* 71/1, 5–19.
- Goodman, Nelson. 1972. "Sense and certainty". In *Problems and Projects*. Indianapolis/New York: Bobbs-Merrill. 60–68.
- Goodman, Nelson. 1983. *Fact, fiction, and forecast*. 4th ed. Cambridge, MA: Harvard University Press.
- Grimm, Stephen. 2010. "The Goal of Explanation". *Studies in the History and Philosophy of Science* 41, 337–344.
- Grimm, Stephen. 2011. "Understanding". In Bernecker, Sven; Pritchard, Duncan (eds). *Routledge Companion to Epistemology*. New York: Routledge, 84–94.
- Grimm, Stephen. 2012. "The Value of Understanding". *Philosophy Compass* 7/2, 103–117.
- Grimm, Stephen. 2016. "Understanding and Transparency". In Grimm, Stephen; Christoph Baumberger; Sabine Ammon (eds). *Explaining Understanding. New Perspectives from epistemology and Philosophy of Science*. New York: Routledge. xx–yy.
- Hills, Alison. 2010. *The Beloved Self. Morality and the Challenge from Egoism*. Oxford: Oxford University Press.
- Hoyningen-Huene, Paul. 2013. *Systematicity. The Nature of Science*. Oxford: Oxford University Press.
- Khalifa, Kareem. 2013. "Is Understanding Explanatory or Objectual". *Synthese* 190, 1153–1171.
- Kuhn, Thomas S. 1977. "Objectivity, Value Judgment, and Theory Choice". In *The Essential Tension. Selected Studies in Scientific Tradition and Change*. Chicago/London: University of Chicago Press. 320–339.
- Kvanvig, Jonathan. 2003. *The Value of Knowledge and the Pursuit of Understanding*. New York: Cambridge University Press.
- Kvanvig, Jonathan. 2009. "The Value of Understanding". In Haddock, Adrian; Millar, Adrian; Pritchard, Duncan (eds). *Epistemic Value*. Oxford: Oxford University Press, 95–111.
- Laudan, Larry. 2004. "The Epistemic, the Cognitive, and the Social". In Machamer, Peter; Wolters, Gereon (eds). *Science, Value, and Objectivity*. Pittsburgh: University of Pittsburgh Press, 14–23.
- Lipton, Peter. 2009. "Understanding without Explanation". In de Regt, Henk W.; Leonelli, Sabine; Eigner, Kai (eds). *Scientific Understanding. Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press. 43–63.

- Pritchard, Duncan. 2010. "Knowledge and Understanding". In Pritchard, Duncan; Millar, Alan; Haddock, Adrian. *The Nature and Value of Knowledge. Three Investigations*. Oxford: Oxford University Press, 1–88.
- Rawls, John. 1999a. *A theory of justice. Revised edition*. Cambridge, MA: Belknap Press.
- Rawls, John. 1999b. "The independence of moral theory". In *Collected Papers*. Cambridge, MA: Harvard University Press. 286–302.
- de Regt, Henk W. 2009. "The Epistemic Value of Understanding". *Philosophy of Science* 76, 585–597.
- de Regt, Henk W. 2014. "Scientific Understanding. Truth or Dare?" In *Synthese*. (DOI 10.1007/s11229-014-0538-7)
- de Regt, Henk W.; Dennis Dieks. 2005. "A Contextual Approach to Scientific Understanding". *Synthese* 144, 137–170.
- de Regt, Henk W.; Leonelli, Sabine; Eigner, Kai (eds). 2009. *Scientific Understanding. Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press.
- Riggs, Wayne. 2003. "Understanding 'Virtue' and the Virtue of Understanding". In DePaul, Michael; Zagzebski, Linda (eds). *Intellectual Virtue*. Oxford: Clarendon Press, 203–226.
- Roche, William. 2012. "Witness Agreement and the Truth-Conduciveness of Coherentist Justification". *Southern Journal of Philosophy* 50, 151–169.
- de Sousa, Ronald. 2010. "Here's How I Feel. Don't Trust Your Feelings!". In Roeser, Sabine (ed.). *Emotions and Risky Technologies*. Dordrecht: Springer. 17–35.
- Stich, Stephen P.; Richard E. Nisbett. 1980. "Justification and the Psychology of Human Reasoning". *Philosophy of Science* 47, 188–202.
- Zagzebski, Linda. 2001. "Recovering Understanding". In Steup, Matthias (ed.). *Knowledge, Truth, and Duty*. New York: Oxford University Press, 235–251.

Notes

- ¹ "Objectual understanding" is due to Kvanvig (2003:191); Pritchard (2010:74) uses "holistic".
- ² See Baumberger et al. 2016 and Grimm 2011 for an overview.
- ³ The classic exposition of explication is Carnap 1962, §§ 2–3. For a detailed discussion see Brun 201*.
- ⁴ We focus on understanding by means of a single theory. Whether and how a subject matter may be understood by means of several, maybe even conflicting, theories cannot be discussed here (but see Elgin 2016).
- ⁵ Elgin (1996:104–105) argues that a theory of justification based on a conception of reflective equilibrium can accommodate non-propositional and even non-verbal elements of theories.
- ⁶ Cf. de Regt 2009:588. Of course, in other contexts we often want to speak of understanding a theory in a less ambitious sense.
- ⁷ We leave open whether the four conditions are collectively sufficient. If it turns out that understanding is incompatible with certain forms of epistemic luck, then it might be necessary to add some safety or sensitivity condition.

- ⁸ Essentialists about biological taxa who think that treating whales as fish was plainly false may consider another example. In the history of economics, introducing a more fine grained periodization may lead to theories which account for more generalizations although they have reduced scope.
- ⁹ Some authors identify epistemic justification with truth conducive justification and argue that justification of theories is not a form of epistemic justification (Laudan 2004); we rather opt for a broader conception of epistemic justification.
- ¹⁰ This is one reason why our proposal is fundamentally different from Kvanvig's (2003; 2009) account of objectual understanding.
- ¹¹ See de Regt 2014 for an argument why such theories can provide understanding without buying the implausible claim that understanding is possible by means of just any old theory from the history of science (e.g. understanding combustion by means of phlogiston theory).
- ¹² The worry that this may result in an account of justification that is too weak is discussed in Sect. 5.2.
- ¹³ Standard accounts of coherence include additional requirements, typically, connectedness and comprehensiveness, which in our account are dealt with as epistemic goals.
- ¹⁴ See Roche 2012 for recent formal work on the details of this objection and what is needed to overcome it.
- ¹⁵ Note that this parallels the division of labour with respect to epistemic justification of knowledge claims: an externalist theory, for example, can argue that justification requires the knower to form her belief through a reliable process, but such theories are silent on which belief-formation processes of which subjects are in fact reliable; for this we need a theory of, e.g. perception.
- ¹⁶ It might be objected that it is natural to say that a theory involves a marginal inconsistency but still has many other virtues and might therefore afford us at least some understanding. We suggest using a strategy of compartmentalization for dealing with this objection (see Elgin 2014): Inconsistencies can be resolved by restricting the theory in a way that eliminates enough of the elements responsible for the inconsistency. But this usually means that the theory fares worse with respect to the epistemic goal of completeness relative to the subject matter, or with respect to simplicity, especially if the restriction must be effected by making exceptions. Thereby, trade-offs which seemingly involve consistency can be modelled as trade-offs with other epistemic goals.
- ¹⁷ Strictly speaking, matters are more complex because the epistemic goals may evolve in the course of the process of equilibrating (cf. Elgin 1996); it may, e.g. become clear that the initially envisaged precision cannot be attained given the intended scope of application.
- ¹⁸ As a consequence, speaking of a *method* of reflective equilibrium may seem misleading since the process is not meant as a recipe that has to be followed. Nonetheless, it may often be a good idea to use such a process as a means for reaching a state of reflective equilibrium.
- ¹⁹ We leave open whether justification by reflective equilibrium must in general – not only in the context of objectual understanding – meet an internalist requirement.

- ²⁰ This is in line with Rawls's (Rawls 1999a:17–8) and Elgin's (1996:128) use of “reflective” as referring to an equilibrium we accept on reflection.
- ²¹ It is not clear that there really are domains in which no position is sufficiently in reflective equilibrium since a reflective equilibrium is relative to the antecedent commitments of some subject. A utilitarian ethics might be sufficiently in reflective equilibrium for somebody with strong consequentialist antecedent commitments. If such a position exists, it also provides an answer to the first version of the worry. This, however, leads to the worry that the account is too weak since any position can be justified relative to weird commitments. We discuss this worry in Sect. 5.2.
- ²² The reason is that condition (b) is a matter of yes or no. Although this is true of condition (a) as well, inconsistent positions are not plausible examples of sufficiently justified positions.
- ²³ Of course, it also helps if the subjects in question accept sufficiently low epistemic standards and content themselves with highly unsystematic theories.
- ²⁴ This point differs from Stich's and Nisbett's proposal of resorting to an expert reflective equilibrium. In expert reflective equilibrium, the gamblers' commitments are simply irrelevant; only the experts' commitments count. In our answer, the gamblers' commitments are still relevant, but we argue that they also have to account for their commitments which result from the testimony of experts.
- ²⁵ In case of some very abstract scientific theories, both abilities may involve the ability to construct suitable models of the subject matter in question (cf. de Regt 2009).