Synthese
DOI 10.1007/s11229-012-0120-0

Conceptual Analysis and Epistemic Progress

Magdalena Balcerak Jackson

Received: 31 August 2011 / Accepted: 30 April 2012

Abstract This essay concerns the question of how we make genuine epistemic progress through conceptual analysis. Our way into this issue will be through consideration of the paradox of analysis. The paradox challenges us to explain how a given statement can make a substantive contribution to our knowledge, even while it purports merely to make explicit what one's grasp of the concept under scrutiny consists in. The paradox is often treated primarily as a semantic puzzle. However, in "Sect. 1" I argue that the paradox raises a more fundamental epistemic problem, and in "Sects. 1 and 2" I argue that semantic proposals—even ones designed to capture the Fregean link between meaning and epistemic significance—fail to resolve that problem. Seeing our way towards a real solution to the paradox requires more than semantics; we also need to understand how the process of analysis can yield justification for accepting a candidate conceptual analysis. I present an account of this process, and explain how it resolves the paradox, in "Sect. 3". I conclude in "Sect. 4" by considering the implications for the present account concerning the goal of conceptual analysis, and by arguing that the apparent scarcity of short and finite illuminating analyses in philosophically interesting cases provides no grounds for pessimism concerning the possibility of philosophical progress through conceptual analysis.

Keywords Conceptual analysis · Paradox of analysis · Epistemic two-dimensional semantics · Philosophical progress

M. Balcerak Jackson (✉)
Emmy Noether Research Group, "Understanding and the A Priori", University of Cologne, Cologne, Germany
e-mail: mbalcerakjackson@me.com

Published online: 05 June 2012

This essay concerns the problem of how we make genuine epistemic progress through conceptual analysis. Many philosophers have rightly become suspicious of the alleged power of our understanding to give us insight into the essences of things. We have also rightly become suspicious of the availability of short and concise statements of necessary and sufficient conditions for the application of many concepts of philosophical interest. Nevertheless, armchair methods continue to play an important role in a great deal of valuable philosophical inquiry. This makes it important to find a defensible, systematic and realistic account of how we use the capacities we retain in the armchair to make substantive discoveries and develop explanatory philosophical theories. To address this general task we need to answer several distinct questions: First, how could it be that our mere grasp of concepts is a source of justification and knowledge for substantive general claims? Second, what is the method by which that source is actually put to work in order to acquire justification? And third, can we see limited reasoners like us as making epistemic progress through conceptual analysis, even if full analyses of concepts rarely seem to be within our cognitive grasp?

Our way into the first two questions will be through consideration of the paradox of analysis. The paradox challenges us to explain how a given statement can make a substantive contribution to our knowledge, even while it purports merely to make explicit what one's grasp of the concept under scrutiny consists in. The paradox is often treated primarily as a semantic puzzle. However, in Sects. 1 and 2 I argue that proposals about the semantic properties of correct analyses—even ones designed to capture the Fregean link between meaning and epistemic significance—are insufficient, on their own, to fully resolve the paradox. Nor, consequently, do they allow us to make much headway on our other leading questions about the method and epistemic value of conceptual analysis as we actually practice is. What we need, both to resolve the paradox and to answer these questions, is an account of how the process of analysis can yield justification for accepting a candidate conceptual analysis, and of the role of our grasp of concepts in that process. I present such an account in Sect. 3, and explain how it resolves the paradox. In Sect. 4 I turn to the implications of the account concerning the goals and value of conceptual analysis, and argue that the apparent scarcity of short and finite illuminating analyses in philosophically interesting cases provides no grounds for pessimism concerning the possibility of philosophical progress through conceptual analysis.

1 The paradox of analysis

The aim of the paradox of analysis is to show that it is impossible for a conceptual analysis to be at the same time both correct and informative.[1] Let us suppose, to begin with, that candidate analyses take the form of statements such as the following:

(1) Humans are rational animals.
(2) Personal identity is unique psychological continuity.
(3) Goodness is what we desire.

[1] The paradox is widely attributed to G. E. Moore, although it arguably reaches back at least to Plato's *Meno* [see Moore (1903) and Plato (1971)].

Clearly, not every true statement of this form qualifies as an analysis. At a minimum, an analysis is meant to hold not only in virtue of contingent correlations between the features mentioned—such as the feature of being human and the feature of being a rational animal in (1)—but to hold, in some sense, as a matter of necessity. Spelling out the relevant sense of necessity is one of the tasks facing an account of conceptual analysis, and we will consider various ways of spelling it out as we discuss responses to the paradox in what follows.

In addition to being (in some sense) necessary, an analysis is meant to capture the meaning of the expression or concept being analyzed. But this latter requirement opens the door to a puzzle. In order to discover or even entertain a correct analysis, the theorist must first understand the expressions or concepts involved. To understand the expressions involved is to know what they mean. But if she already knows their meanings then the analysis, if it is correct, should be uninformative for her. If (1) is the correct analysis of the concept *human*, for example, then it captures the meaning of the concept. But the analyst who grasps the concept *human* already knows what it means. Hence it shouldn't be any genuine advance in her knowledge to conclude, on the basis of conceptual analysis, that humans are rational animals. As Richard Fumerton asks, "…if [the analyst] already knows the meaning of the word or sentence, concept or proposition, how can his work be significant or difficult? How can he know what he means by saying, for example, that this is good or that one thing causes another without knowing what the correct analyses of goodness and causation are?"[2] Conversely, if the analysis *is* informative, if learning it *does* advance one's knowledge, then how can we regard the process that leads to the discovery as merely showing us the meaning of the concept analyzed? This is something already known at the start of inquiry. The puzzle, then, is how anything could qualify as a correct analysis—say, of personhood or goodness or causation—and at the same time make a substantial contribution to our knowledge about those subjects. This is the paradox of analysis.

Sometimes the paradox is treated as primarily a semantic puzzle that arises from certain assumptions about linguistic meaning. For example, suppose that (1) is a correct analysis. Then it would seem that 'human' and 'rational animal' must have the same semantic content; how else could (1) capture the meaning of the concept expressed by 'human'? But if the two expressions have the same semantic content, then (1) should be synonymous with the statement we get by replacing 'human' in (1) with 'rational animal':

(4) Rational animals are rational animals.

Clearly (4) is entirely trivial. If (1) were a correct analysis then it would be no more informative than (4). Since intuitively (1) *is* informative, it follows that it cannot mean the same thing as (4). Nor, therefore, can it be a correct analysis of the concept expressed by 'human'. More generally, the conclusion is that any statement of the relevant form that succeeds in being informative cannot have succeeded in capturing the meaning of the concept in question.

Both this line of reasoning and the previous one lead to the same conclusion, that no conceptual analysis can be both correct and informative. But it is important to

[2] Fumerton (1983, p. 479).

recognize that two significantly different problems hide behind this common conclusion. Call them the *epistemic problem* and the *semantic problem* of analysis. The semantic problem is a puzzle about the semantic features of expressions like 'human' and 'rational animal' and of sentences that contain them. It relies on the assumption that when two expressions are related as *analysandum* and *analysans* then they have the same semantic content, and that they are thus interchangeable in linguistic contexts such as the one occupied by 'human' in (1). It also relies on the assumption, due to Frege, that an uninformative statement like (4) must differ in meaning from an informative statement like (1). The puzzle is about how an analysis can correctly capture the meaning of an expression or concept without being synonymous with a trivial claim like (4). By contrast, the epistemic problem is a puzzle about the epistemic features of the process of conceptual analysis, and of the subjects engaging in that process. It relies on the assumption that there is an intimate connection between understanding a concept and knowing its meaning. It also relies on the assumption that conceptual analysis is a process whereby one utilizes one's understanding of concepts in order to discover the truth of a statement that gives the meaning of the concept under scrutiny. The puzzle is about how this can be a substantive epistemic achievement for someone who already understands the relevant concepts.

It is the epistemic problem that is more fundamental, the problem that needs to be resolved if we are to address the questions about the value of conceptual analysis with which we began. One reason is that it is constitutive of the notion of conceptual analysis that it is a *process*, something that subjects *do*. The epistemic problem is a puzzle about that process, and its solution demands an informative explanation of what we are doing (or ought to be doing) when we engage in it. The semantic problem, by contrast, concerns the static properties of certain sentences that we sometimes come to endorse as the result of engaging in the process of analysis. Moreover, the motivation to engage in conceptual analysis is epistemic: we analyze concepts in order to learn something. The epistemic problem asks, in effect, how such a thing is possible, and no defense of conceptual analysis is complete unless it can answer this question.

The distinction between the semantic and the epistemic problem is not usually made explicit in discussions of the paradox of analysis. One might suspect that this is because the distinction is ultimately not very significant. The semantic problem threatens to show that any correct analysis must have the same meaning as its trivial counterpart, and if this were correct then it would be impossible for the former to be any more informative than the latter. By solving the semantic problem we remove this threat of impossibility, and create room for the possibility that a correct analysis can be substantive discovery; don't we thereby solve the epistemic problem as well? We do not. Showing *that* it is possible for a correct analysis to be informative is one thing; explaining *how* it is possible is another, and this is what we need for a solution to the epistemic problem.To see this point it helps to look at a specific semantic proposal in some detail.

One common approach to solving the semantic problem relies on an appeal to the idea of structured meanings. One version of this approach is developed by Jeffrey King.[3] On King's view the meaning of a syntactically simple predicate like 'brother' is simply the property of being a brother. In contrast, the meaning of a syntactically

[3] King (1998). See also Frege (1950) and Fine (2003a,b).

complex expression is a structured entity whose constituents are the meanings of the syntactic constituents of the expression. So, for example, the meaning of 'male sibling' is a structured entity composed of the meanings of 'male' and 'sibling', i.e. the properties of being male and being a sibling.[4] Because of this difference in structure, 'brother' and 'male sibling' are not strictly synonymous. Nor, consequently, are they freely inter-substitutable with each other, since replacing 'brother' with 'male sibling' yields a sentence whose meaning has a different structure than the original.

(5) Brothers are male siblings.
(6) Male siblings are male siblings.

So on King's view (5) and (6) differ in meaning, just as the reasoning that generates the semantic problem says they should. Nevertheless, according to King this does not prevent (5) from being a correct analysis of 'brother'. The reason is that the property of being a brother is itself constructed out of the properties of being male and being a sibling. This is a claim about metaphysical rather than semantic structure. But because of this isomorphism between the property and the meaning of 'male sibling', the latter succeeds in *representing* the former in a sense that King defines. It is this relation, rather than synonymy, that matters for the correctness of an analysis. Hence statement (5) counts as a correct analysis because the meaning of the complex expression on the right-hand side, by virtue of its semantic structure, successfully represents the property expressed by the term on the left-hand side.[5]

It is important to see that these semantic (and metaphysical) claims do not by themselves shed much light on the epistemic problem of analysis. By appealing to a notion of semantic structure, the proposal makes available fine-grained differences in meaning that allow us to differentiate between statements like (5) and (6). King then avoids the semantic problem of analysis by rejecting its assumption that a correct *analysandum* must have the same meaning as its *analysans*. If this is right, then one can consistently maintain that a statement of the form 'Cs are Fs' is a correct analysis, even though it differs in meaning from the corresponding tautology of the form 'Fs are Fs'. But the mere fact that a correct analysis does not have the latter form does extremely little, by itself, to explain how the discovery of the correct analysis can be a substantive epistemic achievement. After all, isn't 'Cs are Fs' supposed to be something that is given to us just by our grasp of the concepts involved?

Sentence (5) already illustrates this point. Compare (5) and (6) with the following pair:

(7) Knowledge is knowledge.
(8) Knowledge is justified true belief with feature X.

[4] On King's view, in fact, the meaning of 'male sibling' has exactly the same structure as the expression itself. That is, the constituents 'male' and 'sibling' stand in a certain relation to each other in the expression 'male sibling', and the properties of being male and being a sibling stand in *this very relation* to each other in the meaning of the expression. This rather unorthodox view of structured meanings is given further discussion and defense in King (1996), but it is not essential to King's treatment of the paradox of analysis.

[5] It is natural to assume that when this condition is met, it is not metaphysically possible for an entity to fall into the extension of 'male sibling' without having the property of being a brother, and vice versa. Hence presumably King would take the sort of necessity had by correct analyses to be metaphysical necessity, although he does not explicitly address this question.

Imagine that (8) expresses a correct analysis of knowledge. Clearly, coming to know (8) would be a substantive epistemic achievement. On King's view, (7) and (8) are semantically related to each other in just the way that (5) and (6) are. And yet for someone who fully grasps the meaning of 'brother', it is hardly a substantive epistemic achievement to come to know that brothers are male siblings. That a correct analysis does not have the 'Fs are Fs' form does not tell us whether and why it takes genuine (often quite difficult!) epistemic work to discover a correct analysis, even for an analyst who already knows the meanings of the expressions.

One might object that the difference in semantic structure between (7) and (8) actually tells us quite a lot. Statement (7) is a trivial logical truth, and it might be tempting to suggest that, as such, we are in a position to come to know that it is true just by virtue of grasping its semantic structure. But (8) is not a purely logical truth, and grasping its semantic structure does not, by itself, suffice to put us in a position to know that it is true; further epistemic work is required. So once we assign the relevant semantic structures to (7) and (8) we can already see why (8) is a substantive claim—or, at least, by distinguishing it from (7) we can see how it is *possible* for (8) to be substantive. Hence (so the objection runs) the epistemic problem is solved after all.

One problem with this objection is that it remains quite mysterious what it is to grasp the semantic structure of a statement, and how grasping it can yield (or fail to yield) knowledge of its truth. A more serious problem, however, is that the difference between trivial truths and potentially informative correct analyses cannot simply be equated with the difference between statements that are, and statements that are not, knowable in virtue of their semantic structure. This is already illustrated by the fact that there is no epistemic problem of analysis concerning 'Brothers are male siblings', even though it differs in semantic structure from 'Brothers are brothers' in just the way (7) differs from (8).

More generally, let us distinguish between a failure of reasoning and a limitation in evidence. A rational subject can acquire at least prima facie justification for believing that all brothers are brothers just by engaging in certain sorts of reasoning. For example, she might reason as follows: First, consider an arbitrary individual, a, and assume that a is a brother; it follows trivially from this assumption that a is a brother, and hence (discharging the assumption) the conditional conclusion that if a is a brother then a is a brother follows directly; since this argument did not depend on the choice of any particular individual, it follows that for any individual whatsoever, if that individual is a brother then it is a brother; or in other words, all brothers are brothers. Each step in this argument is rational and epistemically appropriate, and notice that it does not utilize any premises for which the subject may have (or lack) evidence. Hence a subject who is able to go through this line of reasoning is already in a position to be justified in believing that all brothers are brothers, even in advance of any evidence that might be taken to support it—evidence such as observed positive instances of brothers who are brothers, a well-confirmed theory that implies that all brothers are brothers, etc. [6] Conversely, if a subject comes to believe that it is not the case that all

[6] One might insist that the subject needs some independent evidence for believing that the rules employed in this argument are valid. Even if this is correct, the present point is that once the subject meets the conditions on the epistemically appropriate use of the relevant inferences—whatever, exactly, those conditions

brothers are brothers, or that it is possible for there to be brothers who are not brothers, then she must have somehow overlooked or ignored or rejected this rational route to the conclusion that all brothers are brothers. In this sense her belief manifests a failure of reasoning.

This is not to say that a subject cannot be justified, all things considered, in withholding belief from the claim that all brothers are brothers. For example, a subject may in fact be capable of going through the reasoning just sketched, but she may have good evidence that the method of conditional proof that it employs, or her ability to utilize it, is not to be trusted. Unless her evidence somehow calls her reasoning into question, however, it cannot defeat the initial justification she has available on the basis of that reasoning.[7]

By contrast, the belief that knowledge is not true, justified belief with feature X need not manifest any failure of reasoning. For example, in the right circumstances a subject might have good evidence that justified true belief is sufficient, by itself, for knowledge, and she might never have encountered the Gettier cases that show that it is not. It is no failure of reasoning to be unaware of Gettier cases; after all, for most of the history of epistemology, no one was aware of such cases. This is merely a reflection of the fact that finite reasoners like us inevitably face limitations in what evidence is available to us. Even in the case of a priori, armchair investigation, we do not already possess all the evidence that might bear on the question we are investigating.

The distinction between failures of reasoning and limitations of evidence allows us to see why the semantic distinctions at issue are inadequate, by themselves, to solve the epistemic problem. Intuitively, someone who believes that not all brothers are male siblings (or vice versa) is just as guilty of a failure of reasoning as someone who believes that not all brothers are brothers. This is so even though it is not something one can come to know just by grasping its semantic structure; failures of reasoning are not limited to mistakes about semantic structure. In the absence of defeaters, one can be in a position to know that brothers are male siblings simply by avoiding such failures, by avoiding any irrational steps when considering the question. By contrast, coming to know (8) requires more than that. It is a matter of overcoming—by luck, or creativity, or methodical determination—the limitations in evidence that might otherwise make it rational to believe some alternative hypothesis, or that might leave one in a state of indecision. It is a matter of gathering the evidence that suffices to warrant us in choosing that hypothesis among the (perhaps) many that are rational options at the start of inquiry. It is in this sense that discovering a correct analysis is a substantive

Footnote 6 continued
are—she is thereby already in a position to have justification for believing that all brothers are brothers. This contrasts, for example, with believing that *a* is G on the basis of one's belief that *a* is F and that all Fs are Gs. Here one requires evidence for these latter beliefs, in addition to whatever evidence may be required for relying on the inference. [Boghossian (2003); Huemer (2002), and Fumerton (2006) provide an entry into the relevant issues concerning the epistemology of inference].

[7] Williamson (2007) imagines subjects who withhold belief from the claim that every vixen is a vixen because of unorthodox views about vagueness or universal quantification. However, these unorthodox views also give them independent grounds for thinking that one or more of the steps in the line of reasoning just sketched is not valid. Hence Williamson's subjects are not cases in which a subject withholds belief from the claim without manifesting a failure of reasoning in the sense intended here. They do manifest a failure of reasoning; they just have good reason to do so.

epistemic achievement. It is a discovery that requires more than merely avoiding failures of reasoning. What we need, for a solution to the epistemic problem of analysis, is an answer to the question of how mere grasp of concepts can take us beyond the mere ability to reason correctly with them, in order to yield substantive discoveries in this sense.

The role of King's semantic proposal, in the preceding discussion, has been to serve as a concrete example to help to illustrate the difference between the semantic and the epistemic problems of analysis. The discussion is not meant to show that there is no way for King himself to solve the epistemic problem. In fact, King recognizes that some statements like (5) count as correct analyses, on his view, despite being trivial and uninformative.[8] He suggests that this is because competence with 'brother' requires a subject to know that the property picked out by 'brother'is composed of the properties of being male and being a sibling. Competence with 'knowledge', on the other hand, requires only that "one's usage of the word be *guided by* the property in the sense that one applies the word to an object iff it possesses the property in question."[9] King offers at least a brief and schematic indication of how competence in this sense might help solve the epistemic problem:

> Because one is able to reliably detect the property in question and so correctly apply the word in question, one can reflect on the circumstances in which the word would correctly apply to an object and reflect on the circumstances in which it would not. By successively dropping features from the circumstances in which the word correctly applies and successively adding features to circumstances in which that word fails to apply, with a bit of insight and ingenuity one might come to see some of the components of the complex property that is the (semantic value) of the word and how they combine to form the complex property. That is, one might be able to assert a sentence that expresses an analysis.[10]

These claims clearly go beyond those that King makes about the semantic differences between correct analyses and their trivial counterparts and, as we will see, they at least tend in the direction of the account developed in Sect. 3. Still, they only just begin to answer the question of how we are able to employ our conceptual capacities to generate justification for substantive philosophical discoveries.[11] And they leave us quite in the dark about whether and why conceptual analysis can yield epistemic progress, even when it falls short of delivering concise and short statements of necessary and sufficient conditions.

[8] King (1998, p. 170).

[9] King (1998, p. 163; italics in original).

[10] King (1998, p. 171).

[11] One challenge King's account would have to answer concerns the notion of being guided by a property, which King claims is sufficient for competence with many expressions. According to the initial characterization of guidance (p. 163), competence with 'knowledge' requires only that one's usage of the term accurately track the actual instantiations of the property of knowledge. In the passage quoted in the main text, however, King suggests that this ability allows competent subjects to determine whether the property *would be* instantiated in various imagined non-actual circumstances. It's not clear why this should be so. I may be able to reliably detect instances of a property via causal interactions, despite having no reliable ability to identify the circumstances under which it would and would not be instantiated upon reflection.

One might suspect that King's semantic proposal explains so little about the epistemology of conceptual analysis because the specific sorts of semantic tools it utilizes are poorly suited for such explanations. As we saw above, the semantic problem arises partly from the Fregean assumption that differences in informativeness correspond to differences in meaning; this is what is supposed to show that an informative analysis 'Cs are Fs' must have a different meaning than the corresponding 'Fs are Fs'. The underlying motivation for the assumption is Frege's view of meaning as epistemic significance (*Erkenntniswert*), the idea that the meanings of statements should play a central role in accounting for such epistemic properties as informativeness, apriority and so on. While King's proposal does provide the required difference in meaning, the purely referential, neo-Russellian meanings it employs do not do a good job of capturing epistemic significance in general.[12] Perhaps a more robustly Fregean theory of meaning would not only be able to solve the semantic problem of analysis, but could do so in a way that would also provide a solution to the epistemic problem.

In my view this is the wrong diagnosis. A semantic proposal—even one that employs an epistemically much richer notion of meaning than King's—is not enough, by itself, to fully solve the paradox of analysis. For a full solution we also need an account of the epistemic basis of conceptual analysis—what are the tools the analyst has by virtue of her grasp of the concepts under scrutiny? And we need an account of the process itself—what are the rational steps that take the analyst from this basis to new insights? To help make the case for this conclusion, in the next section we examine the epistemic two-dimensional semantic framework of David Chalmers and Frank Jackson. While they do not address the paradox of analysis themselves, their semantic framework is arguably the most fully-developed extant attempt to articulate the Fregean connection between meaning and epistemic significance. As we will see, however, even their richer, epistemically-informed semantic values do not suffice to resolve the epistemic problem.

2 Epistemic two-dimensional semantics

Chalmers and Jackson develop epistemic two-dimensional semantics as a framework for specifying semantic contents for concepts, where concepts are thought of as mental representations that contribute to the individuation of representational mental states such as beliefs, desires, perceptions and so on.[13] At the core of the framework are two inter-related assumptions:

[12] The neo-Russellian framework within which King works is premised partly on the conviction that an adequate philosophical account of meaning should eschew the Fregean link between meaning and epistemic significance. [See e.g. Soames (2005); Salmon (1986); Braun (2002).] In fact, an anonymous referee points out that there is a kind of underlying tension here: if differences in epistemic significance need not reflect semantic differences, then it is not clear why the semantic paradox ought to be resolved by means of a semantic proposal like King's in the first place.

[13] Chalmers and Jackson (2001), Chalmers (1996, 2004, 2009); Jackson (1998). For the sake of uniformity I use Chalmers's terminology where it differs from Jackson's (or Chalmers's and Jackson's). There are some differences between Chalmers's and Jackson's articulations of the framework, but these will not matter for our purposes.

(ES1) **Epistemic intensions:** The contents of concepts include at least one component that is determined internalistically and that is bound to epistemic significance.
(ES2) **Modality-apriority link:** There is a specific kind of modality grounded in the epistemic notion of apriority: what is *epistemically possible* is what cannot be ruled out on a priori reasoning, and what is *epistemically necessary* is knowable on a priori reasoning.

Both of these assumptions require further elaboration.

(ES1) posits an aspect or dimension of meaning possessed by concepts. According to Chalmers and Jackson, this aspect of meaning is best captured by a version of possible worlds semantics that assigns a primary intension to each concept. Formally, the primary intension of a concept is simply a function from indices to extensions. Within the epistemic two-dimensional framework, the indices in terms of which the primary intension is characterized are *scenarios*, which correspond to particular fully spelled-out hypotheses about ways the world might be for all we are able to know a priori. We can think of scenarios as possible worlds of a certain sort. However, it is important to keep in mind that scenarios are possible worlds that are characterized epistemically: a scenario is a complete way for the world to be that cannot be ruled out a priori. For example, empirical investigation has revealed that caffeine—the stimulant present in coffee and tea—is actually the chemical substance 1,3,7-trimethylxanthine. Given this discovery, familiar arguments from Kripke and Putnam make it plausible that it is not metaphysically possible for caffeine to be anything other than 1,3,7-trimethylxanthine. However, the discovery is presumably not something that we could come to know a priori, and hence there are scenarios at which caffeine is not 1,3,7-trimethylxanthine. We can imagine discovering that the substance that is actually contained in coffee and tea, and that is responsible for their stimulant effects, turns out to be, say, 2,4,8-paraxanthine instead of 1,3,7-trimethylxanthine. Under this hypothesis it is rational to conclude that caffeine is 2,4,8-paraxanthine, and so the primary intension of *caffeine* picks out 2,4,8-paraxanthine at such a scenario. The primary intension thus varies from scenario to scenario, depending on the various discoveries about the underlying nature of caffeine that are compatible with what we can know a priori. The primary intension of a concept is meant to reflect the way in which its meaning puts rational constraints on which entities fall into its extension, given various different hypothetical assumptions about how things turn out to be in the actual world.

The primary intension of *caffeine* contrasts with its *secondary intension*. Secondary intensions are defined in terms of metaphysical, counterfactual possibilities—ways things could have been—rather than epistemic possibilities—ways things might actually be from a certain epistemic standpoint. According to Kripke and Putnam, caffeine could not have been anything other than 1,3,7-trimethylxanthine; if coffee had contained 2,4,8-paraxanthine instead of 1,3,7-trimethylxanthine then coffee would not have contained caffeine. Hence the secondary intension of *caffeine*, unlike its primary intension, picks out 1,3,7-trimethylxanthine at every metaphysically possible world. For just this reason, secondary intensions do not do a very good job of capturing the epistemic significance of the concept *caffeine*. It is by virtue of positing distinct primary and secondary intensions as two aspects of the meanings of concepts that the

Chalmers and Jackson framework earns the label "two- dimensional" semantics. But it is the primary intensions that do the work of articulating the Fregean link between meaning and epistemic significance.

Let us turn to (ES2). (ES2) posits a kind of necessity and possibility that is characterizable in epistemic terms. Part of the import of positing this epistemic modality is that a corresponding space of possible worlds can be extracted from it in the familiar way, and it is this space of possible worlds—scenarios—that is used to characterize the dimension of meaning posited in (ES1).[14] (ES2) also provides a promising way of specifying the sense in which correct conceptual analyses are necessary, the need for which was noted in Sect. 1: the sort of necessity had by a correct analysis is epistemic necessity, which is represented within the two-dimensional framework as having a primary intension that yields truth at all scenarios. Moreover, according to (ES2) this is a sort of necessity that is, in principle, open to investigation solely on the basis of a priori rational reflection. If Kripke and Putnam are right then at least some of the facts about what is metaphysically necessary and possible are only knowable on the basis of empirical investigation. But this is not so for facts about what is necessary and possible in the epistemic sense posited in (ES2). In addition, since the epistemic possibilities posited by (ES2) are used to characterize primary intensions, the framework grounds the possibility of acquiring a priori knowledge by investigating primary intensions.

Both (ES1) and (ES2), and the two-dimensional semantic framework they ground, depend on a highly idealized notion of apriority. This idealized notion can be seen at work in the characterization of epistemic modality in (ES2): the epistemic possibilities it posits are possibilities that cannot, in principle, be ruled out a priori—that is, not even by an ideal reasoner with full grasp of the necessary concepts, who is unlimited and infallible in her powers of reasoning, memory, attention and other cognitive capacities. The characterization of primary intensions involves a corresponding idealization: the extension of a concept at a scenario reflects the judgments of an ideal reasoner about what falls under the concept given the hypothesis that the scenario obtains.[15] Primary intensions thus reflect the a priori inferential structure of concepts for an ideal reasoner, rather than for limited and fallible reasoners like us.

Certainly, one can have doubts about how closely ordinary subjects approximate Chalmers's and Jackson's rational ideal in the way just sketched.[16] More generally,

[14] Chalmers typically takes the space of scenarios to constitute a second modal space distinct from that of the metaphysically possible worlds. Jackson takes epistemic and metaphysical modality to correspond to two ways of evaluating statements at points in a single modal space. This difference leads to important differences in their interpretations of the two-dimensional semantic framework, but for our purposes it is a difference we can safely ignore.

[15] Hypotheses in the sense needed here are fine-grained: the hypothesis that Hesperus rises in the north is distinct from the hypothesis that Phosphorus rises in the north, even though Hesperus is identical with Phosphorus. The question therefore arises of how the hypotheses that determine primary intensions are to be specified. According to Chalmers and Jackson (2001) they should be given in a restricted, canonical vocabulary that contains only terms from microphysics, terms for describing the phenomenal qualities of experiences, and indexical terms. This suggestion amounts to the hypothesis that knowing what the world is like in microphysical, phenomenal and indexical respects puts an ideal subject in a position to infer everything else about what the world is like *a priori*. The hypothesis is disputed by Block and Stalnaker (1999).

[16] See Schroeter (2004) for one such challenge.

one can raise various sorts of challenges to the epistemic two-dimensional semantic framework.[17] I think these challenges can be answered. However, our current question is not whether the semantic claims of the framework can be defended, but whether, if they are correct, they provide a solution to the epistemic problem of analysis. And the answer is that they do not.

To see why, let us return to our imagined correct analysis of the concept *knowledge*:

(8) Knowledge is true, justified belief with feature X.

If (8) is a correct analysis then presumably its truth is accessible to an ideal reasoner with a perfect grasp of the concept knowledge and the other concepts involved, and with unlimited and infallible resources for reasoning. Hence (8) is epistemically necessary; indeed, as we just saw, within the epistemic two-dimensional framework it is plausible to regard epistemic necessity as a necessary condition on the adequacy of an analysis. But surely (7), repeated below, is epistemically necessary as well.

(7) Knowledge is knowledge.

Within the epistemic two-dimensional framework, both (7) and (8) are true at every scenario. Or to put it another way, for any scenario whatsoever, an ideal reasoner is in a position to figure out that both (7) and (8) are true given that the scenario obtains. It follows that (7) and (8) have the same primary intension and hence, within the epistemic two-dimensional framework, the same meaning. Yet (8) is clearly informative in a way that (7) is not; discovering a correct analysis of the form of (8) would be an incredible achievement in the study of knowledge. Therefore we are faced with a version of the semantic problem of analysis: for (8) to qualify as a correct analysis it must have the same meaning as (7), but to respect Fregean intuitions about epistemic significance the two statements must have different meanings. The epistemic two-dimensional semantic framework, as developed so far, does not solve this problem.

This result should not be very surprising. It is really just a special instance of a more familiar general problem for possible worlds semantics. The problem is that sets of possible worlds (or the primary intensions of statements, which are the characteristic functions of such sets) are too coarse-grained to provide the intuitive differences in meaning we find between many modally equivalent statements. For example, any two statements that are knowable a priori are true at every scenario, and hence all a priori statements have the same meaning on this account. And primary intensions cannot distinguish between the members of a priori equivalent pairs such as 'The number of Obama's children is two' and 'The number of Obama's children is the only even prime number'. The failure to distinguish between correct informative analyses and their tautological counterparts is a further illustration of the fact that primary intensions are too coarse-grained to capture the intuitive distinctions between meanings we recognize.

This familiar problem has a familiar solution, which is to appeal to semantic structure in order to make finer-grained distinctions among meanings. Hence Chalmers

[17] See, for example, Schroeter (2006) and Soames (2005).

introduces the notion of a structured primary intension of a statement: an entity whose constituents are the primary intensions of the basic expressions contained in the statement, and which is constructed in a way determined by the statement's logical form.[18] Statements (7) and (8) differ in their logical form, and they contain expressions with different primary intensions. Hence they have different structured primary intensions.

In this way the appeal to structured primary intensions yields a solution to the semantic problem of analysis. In fact, the solution exactly parallels the one from King that we examined in Sect. 1, but it is developed within a semantic framework that is designed to capture the Fregean link between meaning and epistemic significance. So does it constitute progress on the epistemic problem of analysis? I think it is clear that it does not. The differences in semantic structure that Chalmers posits do not, by themselves, shed any light on why coming to know (8) is a substantive epistemic achievement in a way that coming to know (7) is not. By hypothesis both statements are epistemically necessary; that is, in both cases it is possible, in principle, to come to know the truth of the statement on the basis of ideal a priori reflection. But only in one case is it a substantive epistemic achievement to do so. The fact that (7) and (8) have different structured primary intensions tells us very little about why this is so. Once again, the appeal to structure provides the differences in meaning required to solve the semantic problem of analysis, but those differences in meaning aren't sufficient to explain the differences in epistemic significance that generate the epistemic problem.

Like King, Chalmers and Jackson at least implicitly acknowledge that the epistemic problem requires us to go beyond the semantic properties of analyses, and say something about the actual process of conceptual analysis:

> When given sufficient information about a hypothetical scenario, subjects are frequently in a position to identify the extension of a given concept, on reflection, under the hypothesis that the scenario in question obtains. Analysis of a concept proceeds at least in part through consideration of a concept's extension within hypothetical scenarios, and noting regularities that emerge.[19]

These remarks, like King's, at least hint at a way to approach the epistemic problem. But developing a full solution requires some positive explanation of why discovering a correct analysis requires more than merely avoiding failures of reasoning, so that it can in this sense be a substantive epistemic achievement, and yet is a discovery that is grounded in one's grasp of the concepts involved. For this we need a better understanding of the *cognitive dynamics* of conceptual analysis—the process by which we are able to utilize our grasp of concepts in order to acquire the evidence that allows us to rationally settle on a correct analysis. In the next section I provide a sketch of the cognitive dynamics of conceptual analysis, and explain how it resolves the epistemic problem.

[18] See Chalmers (2011).

[19] Chalmers and Jackson (2001, p. 322).

3 The cognitive dynamics of conceptual analysis

It is a familiar idea that we can think of a candidate analysis as analogous to a scientific hypothesis, and that we should see conceptual analysis as a process in which such hypotheses are subjected to testing and evaluation. This idea can be made more precise, and developed into an account of the cognitive dynamics of conceptual analysis, by answering the following questions:

(a) What sort of hypothesis is a candidate analysis?
(b) What is the nature of the evaluation to which such a hypothesis is subjected?

The central component of the process of conceptual analysis is the use of thought experiments—the consideration of hypothetical or imaginary cases in order to elicit judgments employing the concepts under scrutiny. Adequate answers to questions (a) and (b) should make it clear how the use of thought experiments contributes to the evaluation of candidate analyses. And if our account is to help solve the epistemic problem of analysis, it should also make it clear how knowledge (or justified belief) reached in this way is grounded in the grasp of concepts.

Let us begin with question (a). Suppose that C is an expression for the *analysandum* under scrutiny, and F is a distinct (perhaps complex) expression for a candidate *analysans*. The hypothesis to be tested via conceptual analysis can be characterized, roughly, as the hypothesis that F yields epistemically necessary and sufficient conditions for the application of C. The epistemic two-dimensional framework provides a useful set of tools for making this suggestion more precise, by allowing us to represent a candidate analysis as an explicit generalization about the space of epistemically possible scenarios:

(9) $\forall s \forall x (\mathrm{C}(x, s) \leftrightarrow \mathrm{F}(x, s))$

Here 's' is a variable ranging over scenarios; a statement of the form of (9) says that for any given scenario, something is an instance of C at that scenario if and only if it is an instance of F at that scenario.[20] For example, the candidate analysis of knowledge as justified true belief can be thought of as the hypothesis that every instance of knowledge at every scenario is an instance of justified true belief at that scenario, and vice versa. It is a claim about the relative distribution of cases of knowledge, and of cases of justified true belief, across epistemic modal space.[21]

In order to answer question (b), we also need to introduce *extensional statements* corresponding to intensional statements of the form in (9). Whereas candidate analyses

[20] Here, for simplicity, we are confining our attention to cases in which C and F take only a single argument over and above the argument for scenarios. The choice to employ explicit quantification over scenarios (rather than a modal operator) is to help make vivid the suggestion that candidate analyses are *generalizations* employing the concept under scrutiny. Note that there is no suggestion here that every true statement of the form given in (9) is an adequate analysis. Nor, for reasons to be discussed in the concluding section, should we think that the goal of conceptual analysis is confined to the discovery of statements of this form.

[21] Recall the need for an account of conceptual analysis to explicate the sense in which a correct analysis is necessary, noted in Sect. 1. On the present proposal, a correct analysis is, in effect, a true claim that such-and-such is epistemically necessary. Whether this claim is itself epistemically (or metaphysically) necessary is a further question that we need not attempt to settle here.

make general claims about the whole space of scenarios, extensional statements make claims about specific individual scenarios. These are simply statements of forms like the following:

(10) $C(i, s^*), F(i, s^*), \neg C(i, s^*), \neg F(i, s^*)$

Here 's^*'is a term picking out a certain scenario and 'i' is a term picking out some individual. Statements of the forms in (10) are not modal claims; rather, they simply say that some individual i is (or is not) an instance of C or F at s^*. For example, corresponding to the candidate analysis of knowledge as justified true belief, there are extensional statements to the effect that a particular subject, say Smith, is (or is not) an instance of knowledge at s^* and is (or is not) an instance of justified true belief at s^*. Whereas candidate analyses make claims about the distributions of the instantiations of properties across modal space, extensional statements concern the instantiation of properties by specific individuals at specific scenarios.

The significance of extensional statements for question (b) is as follows. We can think of the sorts of hypothetical cases employed in thought experiments as sketches of epistemically possible scenarios. And we can think of our judgments about those cases as informing us about the truth or falsehood of various extensional statements of the forms just sketched. For example, we can consider a hypothetical case in which a given subject, Smith, believes that Jones owns a Ford and, on that basis, infers that either Jones drives a Ford or Brown is in Barcelona. We can suppose that although Jones doesn't drive a Ford, Smith has very good reason to believe that he does, and he has no independent reason to believe that Brown is in Barcelona. We can think of this hypothetical case as describing a certain scenario s, and by considering the case we come to learn the truth of the extensional statement that Smith does not have knowledge at s. We also come to learn the truth of the extensional statement that Smith has justified, true belief at s. These extensional statements then become evidence that we can use to evaluate various hypotheses about the distribution of cases of knowledge across the space of scenarios. The relationship between a candidate analysis and the corresponding extensional statements is thus the relationship between hypothesis and data. The primary purpose of conducting thought experiments is to gather data relevant for evaluating the hypothesis under scrutiny.

Representing thought-experimental hypothetical cases via epistemically possible scenarios, in the technical sense introduced in the previous section, involves a bit of an idealization. When considering a hypothetical case such as Gettier's, for example, we are not really evaluating a full scenario in Jackson's and Chalmers's sense, but rather a small "cut-out" that leaves many things open about what the rest of the world is like. In effect, to judge that the subject in the hypothetical Gettier situation does not have knowledge is to commit oneself to the same judgment about a whole range of scenarios that differ in ways left open by the description of the case. This commitment can turn out to be mistaken, if it turns out that some way in which the scenarios differ matters for whether the subject has knowledge or not. Skilled conceptual analysts try to be attuned to such possibilities and take steps to rule them out, just as skilled empirical investigators try to identify and control for potentially relevant variables. But there is no guarantee of success. Certainty is no more attainable in the case of conceptual analysis than in most other (perhaps any) human epistemic endeavors.

How do we use extensional statements to evaluate candidate analyses? One way in which extensional statements can bear on an analysis is no doubt already clear from the example just given. Consider the classical analysis of knowledge as justified true belief. This analysis is the hypothesis that all and only cases of knowledge at any given scenario are cases of justified true belief at that scenario. The extensional statements about Smith in the scenario just described together provide a straight forward counter-instance to this generalization, and hence they provide very strong evidence that the hypothesis is false. The consideration of such *refuting scenarios* is one straightforward way in which extensional statements bear on candidate analyses.

But thought experiments do not merely refute candidate analyses. On the contrary, the data gathered from thought experiments also has the potential to provide positive support for a candidate analysis. For example, there is a wide and diverse range of hypothetical (and actual) cases of justified true beliefs that are not "lucky" or accidental in the way the case involving Smith is, and these are cases of knowledge. The classical analysis offers a simple and elegant explanation of this pattern, and so these hypothetical cases provide positive support for accepting the classical analysis. It is only once Gettier cases are discovered that this support gets overridden; this is why there was no failure of reasoning involved in accepting the classical analysis of knowledge at an earlier stage of inquiry. Even refuting scenarios provide positive epistemic support, and not just reason to reject an analysis. Discovering systematic patterns among counterexamples, and understanding why they are counterexamples, can improve our insight into the concept under scrutiny. It can suggest new directions for analysis, new ways to adjust our hypothesis so as to better explain the range of hypothetical cases we observe.

The picture here of how individual cases provide positive support for general hypotheses is one that is already familiar from the empirical sciences. The relevant form of reasoning is inference to the best explanation. What we are doing when we propose an analysis for a concept C is forming a general hypothesis about the features of scenarios that account for those scenarios being, or failing to be, instances of C. We collect evidence for or against this hypothesis by engaging in thought experiments and noticing the patterns among extensional statements that emerge. Throughout this process, we aim to design appropriate thought experiments in very much the way empirical scientists design appropriate experimental settings, in order to test the predictions of the hypothesis, control for potentially relevant factors, and so on. We have confirmation for our candidate analysis insofar as it does a better job than its rivals of explaining the evidence we gather.[22] Inference to the best explanation is at work even in the use of refuting scenarios. When faced with an apparent counterexample, we always have the option of preserving our analysis by explaining it away as somehow spurious. To treat the counterexample as genuine is to reject such an explanation as not good enough, it is to judge that a more promising explanatory strategy would be to revise our analysis in light of it.[23] Hence inference to

[22] Here, as in the empirical sciences, the goodness of an explanation is determined according to familiar, although hard to articulate criteria of precision, simplicity, unification, and so on. See Lipton (2000) for a survey of the epistemological challenges facing accounts of inference to the best explanation.

[23] According to Weatherson (2003), the explanatory virtues of the classical analysis of knowledge so strongly outweigh the counter-evidence of Gettier cases that we ought to regard them as spurious even if we cannot currently explain them away.

the best explanation is always at the core of the cognitive dynamics of conceptual analysis.[24]

One might worry that since inference to the best explanation is an inductive, uncertain method, it cannot provide sufficient justification for claims about epistemically necessary connections among concepts.[25] However, according to the present picture, the inductive base includes individuals and scenarios that are potentially distributed across the full range of epistemically possible worlds. I see no reason, in principle, why such a base could not provide good inductive grounds for accepting a hypothesis about what is epistemically necessary. Of course, since inference to the best explanation is inductive, it cannot provide certainty for its conclusions. But certainty is not to be had in conceptual analysis any more than in other areas of inquiry.

We can illustrate the general pattern of reasoning involved in the process of analysis by means of a (somewhat schematic) sketch of the attempt to analyze personal identity. Suppose we begin with the *prima facie* plausible hypothesis that the diachronic persistence of a person depends on the persistence of her body.[26] According to this *Body Hypothesis*, a person $p_1$at time $t_1$is identical with person $p_2$at time $t_2$at a given scenario if and only if there is a sufficient degree of bodily continuity between $p_1$at $t_1$and $p_2$at $t_2$at that scenario. We then test this hypothesis by considering various scenarios—involving surgery to replace parts of p_1's body with parts of someone else's body, accidents that involve massive damage to p_1's body, and so on—and asking ourselves in each case whether the relevant extensional statements about p_1's survival predicted by the Body Hypothesis are true. The result of this inquiry is that the Body Hypothesis is probably false, and that a much better explanation of the facts about p_1's survival at the various scenarios is that the persistence of a person requires only the persistence of certain psychological features and relations. This gives us epistemic support for a new *Psychology Hypothesis*: a person $p_1$at time $t_1$is identical with person $p_2$at time $t_2$at a given scenario if and only if there is a sufficient degree of psychological continuity between $p_1$at t_1 and p_2 at $t_2$at that scenario. However, as Derek Parfit discovered, there are scenarios in which more than one person fulfills the criteria at $t_2$for being psychologically continuous with $p_1$at t_1.[27] In such cases, the Psychology Hypothesis collides with our tacit assumption that personal identity is a transitive relation. We are then faced with a choice of how to respond to Parfit's data. Do we revise our assumption about the transitivity of personal identity? Do we look for some further condition with which to supplement the Psychology Hypothesis? Or do we abandon psychological continuity and look for an analysis in entirely different terms? In reality, of course, all of these strategies can be pursued at once. The point is that, as in the earlier steps in the process of analysis, the results of these strategies

[24] Goldman and Pust (1998) suggest that thought experiments yield intuitions about single cases that serve as evidence in something like the way suggested here. However, Goldman and Pust see such evidence as primarily confirming psychological generalizations about mental representations, rather than as confirming epistemic necessities involving the concepts under scrutiny.

[25] Thanks to an anonymous referee for this objection.

[26] Ayer (1936).

[27] Parfit (1984).

are to be weighed according to how well they explain the facts about the persistence of persons under various sorts of epistemically possible conditions.

Finally, it remains to ask: what role is played by the grasp of concepts in the process of analysis? On the present proposal, the grasp of concepts comes into the picture at the very beginning of the process—at the stage in which we gather the extensional statements that constitute the evidence against which candidate analyses are to be weighed. It is by virtue of grasping the concepts under scrutiny that we are in a position to know the truth-values of the relevant extensional statements. This assumption amounts to a moderate form of rationalism: rational subjects have conceptual intuitions about hypothetical cases, and these intuitions yield reliable judgments about the truth and falsehood of extensional statements employing (expressions for) concepts that they grasp. Of course, this is not to say that we have an infallible ability to correctly judge extensional statements in all cases. Just as empirical scientists who are competent with the relevant laboratory apparatus can reliably—but not infallibly—determine, say, the presence or absence of some pertinent molecule in samples of a substance, conceptual analysts with a competent grasp of the concepts under scrutiny can reliably determine, say, the presence or absence of knowledge in particular scenarios. It is the application of this ability, in conjunction with the application of general reasoning abilities to evaluate competing explanations, identify further relevant scenarios and so on, that ultimately yields justification for accepting or rejecting a candidate analysis. It is in this sense that the process of conceptual analysis is grounded in our grasp of concepts.[28]

It is hopefully clear by now how the cognitive dynamics just sketched helps us to resolve the epistemic problem of analysis. Earlier I argued that resolving the problem requires us to make it clear how the discovery of a correct analysis can be grounded in our grasp of concepts, and yet can be more than just a matter of avoiding failures of reasoning employing those concepts. According to the account given here, the justification we have for accepting a candidate analysis in any given case is based on the evidence we have collected via the systematic use of thought experiments. But there are always limits in the evidence available to us, simply because finite creatures like us do not have the capacity to survey the whole space of epistemically possible scenarios. It can very well happen that the limited evidence available to us makes it most reasonable, all things considered, to accept a candidate analysis that is not correct. Let us return to the pair (7) and (8):

(7) Knowledge is knowledge.
(8) Knowledge is justified true belief with feature X.

As we saw above, one has justification for accepting (7) just on the basis of appropriate reasoning, even in the absence of any specific evidence. Not so for (8). Depending on what scenarios one has considered, one might not have encountered the evidence that

[28] What would explain competent subjects' reliable capacity to judge extensional statements? On the Chalmers and Jackson account, the primary intension of a concept is determined by the *a priori* judgments that an ideal reasoned would make about its extension in each scenario. To the extent that ordinary reasoners like us at least approximate this ideal, our judgments about extensional statements will tend to match those of the ideal reasoned, and so will tend to be correct. Other explanations are possible. [See, for example, Peacocke (2003, 2004).]

would show that justified true belief is not sufficient for knowledge, and it might be that the scenarios one has considered are best captured by the classical analysis. If so, then it would be no failure of reasoning to reject (8) in favor of the classical analysis. Competent reasoning with the concept *knowledge* does not automatically require one to accept (8); one must have evidence that makes it rational to rule out salient alternatives. This is why finding a correct analysis is a substantive discovery—it requires the accumulation of good evidence, by reflection on a wide range of diverse scenarios, in addition to reasoning appropriately.

Moreover, on the account developed here this is so even though the evidence is derived from one's grasp of concepts. Recall that Fumerton poses the paradox of analysis by asking how the analyst can know the meaning of, say, 'goodness' or 'causation' without already knowing the correct analyses of the concepts expressed. On the present account, knowing the meaning of 'goodness' only puts one in a position to know, of various individual cases presented in certain ways, whether or not they are cases of goodness. This falls far short of already having knowledge of the correct analysis; the analyst need not even have any justification for believing it at the outset. She needs to put her grasp of the concept to work, gather the necessary evidence and use it to correctly evaluate various competing hypotheses. But if and when she comes to know the correct analysis of goodness, this knowledge is grounded in her grasp of the relevant concepts. Thus the account developed here answers the challenge posed by the epistemic problem of analysis, to explain how the discovery of a correct analysis can be grounded in our grasp of concepts, and yet can nevertheless be a substantive epistemic achievement.

Many of the basic elements of this answer to the epistemic problem of analysis are familiar. The analogy between thought experiments and scientific data-gathering is, of course, already encouraged by the label 'thought experiment'.[29] And as we have seen, both King and Chalmers and Jackson make the suggestion, at least in passing, that conceptual analysis should somehow be seen as an abductive enterprise.[30] But as developed here, these suggestions have some consequences that are not often recognized. One such consequence is that there need not be any sense in which the analyst already knows or believes the central conceptual truths concerning the concept under scrutiny, prior to coming to know those truths via successful conceptual analysis. As we have seen, a subject with a fully competent grasp of the concept *knowledge* may justifiably hold false beliefs about the epistemically necessary or sufficient conditions for knowledge. Such a subject might not even be disposed to accept the correct analysis, except in the extremely weak sense that she might come to discover it by a careful and fortuitous application of her conceptual capacities.[31] It is often alleged that there

[29] See Gendler (2000) and Sorensen (1992) for classic discussions of the analogy between scientific experiments and thought experiments.

[30] Ludwig (2007) also suggests that the abductive nature of conceptual analysis helps resolve the paradox of analysis. But Ludwig's main focus is on challenges to conceptual analysis from experimental research on responses to thought experimental scenarios.

[31] Ludwig (2007) characterizes subjects as having implicit knowledge of a correct analysis, by virtue of having certain abilities to apply the concepts employed in the analysis. But this characterization is misleading, at best. It is no more appropriate than to describe someone as having implicit knowledge of all the truths of arithmetic by virtue of being able to do arithmetical calculations, or to describe someone as having

are no truths concerning knowledge that a subject must be disposed to accept, in order to count as grasping the concept *knowledge*. This is often taken to be a serious threat to the viability of conceptual analysis as a philosophical method. But this possibility fits quite comfortably within the picture of conceptual analysis developed here.[32]

Another consequence is that the project of developing a neo-Fregean theory of meaning within the epistemic two-dimensional semantic framework faces an inevitable limitation. One goal of the two-dimensional framework, as Chalmers and Jackson use it, is to capture the Fregean link between meaning and epistemic significance. For a scenario H to be in the primary intension of a statement S is for the hypothesis that H obtains to make it rational, under ideal *a priori* reflection, to accept S. This approach treats differences in epistemic significance as differences in which scenarios verify the statements at issue under ideal *a priori* reflection. On the present picture, however, to account for differences in epistemic significance between a correct analysis like (8) and its trivial counterpart (7), we need to pay attention to differences in the way the statements can come to be known by ordinary, limited subjects like ourselves. Moreover, we need to consider how distributions of and patterns among facts *across* scenarios can provide such subjects with rational support a statement; this is not just a matter of determining which scenarios verify the statement and which do not. Hence the epistemic features of statements that account for the relevant differences in significance do not correspond to any of the sorts of semantic values definable within the two-dimensional framework. There is a central aspect of epistemic significance that the Chalmers and Jackson theory of meaning cannot capture.

4 Epistemic progress and the goal of analysis

In Sects. 1 and 2 we focused on the first of the three question with which we began: how is it possible to make substantive philosophical discoveries, just by relying on our grasp of concepts? The upshot of the discussion in those sections was that an adequate solution requires attention to the second of our three leading questions, the question of what method of inquiry actually enables us to employ our conceptual capacities as a source of justification for candidate analyses. Developing an answer to this question, and showing how it resolves the epistemic problem, was the task of Sect. 3. In light of this discussion, we can now conclude by returning to the last of our three leading questions: How are we to make sense of the evident fact that conceptual analysis so rarely seems to yield concise, complete analyses of the concepts in which we are interested as philosophers? What is the value of conceptual analysis, if it never reaches its goal?

It is a consequence of the present view that the epistemic value of conceptual analysis in no way depends on whether or not it is possible to find finite and short informative

Footnote 31 continued
knowledge of all the optimal moves in every chess game by virtue of being able to play chess. Coming to know a correct analysis requires acquiring good evidence, and prior to the process of analysis, there is no useful sense in which we already have that evidence.

[32] The most recent sustained challenge of this sort can be found in Williamson (2007). A response to Williamson's challenge that makes use of the picture of conceptual analysis given here is offered in Balcerak and Balcerak (2011).

statements of epistemically necessary and sufficient conditions for most, or even any, philosophically interesting concepts. Consider the personal identity debate sketched in the previous section. Perhaps this debate has not managed to yield a full analysis of what it is for a person at one time to be identical with a person at some other time; perhaps it will never yield such a thing. Nevertheless, even if this is the case, conceptual analysis has given us substantial philosophical insight into personal identity. We are in a position to explain quite a lot about the distribution of cases of personal identity across epistemic modal space, even in the absence of a full classical analysis. And we have at least a partial understanding of how that distribution relates to other salient features, such as the relations of bodily, organismic and psychological continuity that obtain or fail to obtain in various scenarios. This represents genuine progress, a genuine increase in our knowledge and understanding of personal identity.

One reason for skepticism about conceptual analysis—a highly influential one among many contemporary philosophers, I suspect—is the history of failed attempts to arrive at full and informative analyses. But partial and approximate analyses are worth having, because they still have explanatory power, and so conceptual analysis is worth doing regardless of whether it ever leads to finite and short illuminating analyses. Doubts about the availability of analyses in the classical sense are no reason for skepticism about conceptual analysis.

The situation here is not so different than in the empirical sciences. For scientists in many domains the goal of a simple, exceptionless and explanatory theory of the target phenomena is a distant ideal at best. Actual scientific practice aims to discover explanations that are expected to turn out to be incomplete or partial in various respects, and the goal of theorizing is to develop better and better approximations of the truth. On the picture of conceptual analysis presented here, the target phenomena are facts about the distribution of instances of knowledge, of relations of personal identity, and so on, across the domain of epistemically possible scenarios. A concise statement of the epistemically necessary and sufficient conditions for something's being an instance of knowledge may be a distant ideal. But that gives us no reason at all not to look for good explanations of why actual or conceivable situations of various sorts count as ones in which the subject knows or fails to know, even if such explanations inevitably turn out to be merely partial or approximate. Both the value and the goal of conceptual analysis, like the value and the goal of empirical science, is the continual pursuit of better and better approximations.

Acknowledgments For helpful comments and discussion I thank Yuri Cath, David Chalmers, Patrick Forber, Thomas Grundmann, Daniel Stoljar, two anonymous referees, audiences at the Conference of the Australasian Association for Philosophy, the Conference of the Gesellschaft für Analytische Philosophie and the Department Colloquium of the University of California at Davis, and especially Brendan Balcerak Jackson. Research on the paper has been generously supported by the DAAD and the DFG.

References

Ayer, A. J. (1936). *Language, truth, and logic*. London: Gollancz.

Balcerak Jackson, M., & Balcerak Jackson, B. (2011). Understanding and philosophical methodology. *Philosophical Studies*. doi:10.1007/s11098-011-9729-y.

Block, N., & Stalnaker, R. (1999). Conceptual analysis, dualism and the explanatory gap. *The Philosophical Review, 108*, 1–46.
Boghossian, P. (2003). Blind reasoning. *Aristotelian Society Supplementary Volume, 77*, 225–248.
Braun, D. M. (2002). Cognitive significance, attitude ascriptions, and ways of believing propositions. *Philosophical Studies, 108*, 65–81.
Chalmers, D. J. (1996). *The conscious mind*. New York: Oxford University Press.
Chalmers, D. J. (2002). Does conceivability entail possibility? In T. Gendler & J. Hawthorne (Eds.), *Conceivability and possibility*. New York: Oxford University Press.
Chalmers, D. J. (2004). Epistemic two-dimensional semantics. *Philosophical Studies, 18*, 153–226.
Chalmers, D. J. (2009). The two-dimensional argument against materialism. In B. McLaughlin (Ed.), *Oxford handbook of the philosophy of mind*. New York: Oxford University Press.
Chalmers, D. J. (2011) Propositions and attitude ascriptions: A fregean account. *Noûs 45*(1), 595–639.
Chalmers, D. J., & Jackson, F. (2001). Conceptual analysis and reductive explanation. *The Philosophical Review, 110*, 315–361.
Fine, K. (2003a). Moore's paradox of analysis. Oxford Lecture.
Fine, K. (2003b). *The limits of abstraction*. Oxford: Oxford University Press.
Frege, G. (1950). *The foundations of arithmetic: A logicomathematical enquiry into the concept of number*. Oxford: Blackwell Publishing (translated by J. L. Austin).
Fumerton, R. A. (1983). The paradox of analysis. *Philosophy and Phenomenological Research, 43*, 477–497.
Fumerton, R. A. (2006). *Epistemology*. Oxford: Blackwell Publishing.
Gendler, T. Só. (2000). *Thought experiments: on the powers and limits of imaginary cases*. New York: Garland.
Goldman, A., & Pust, J. (1998). Philosophical theory and intuitional evidence. In M. R. DePaul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*. Lanham: Rowman & Littlefield.
Huemer, M. (2002). Fumerton's principle of inferential justification. *Journal of Philosophical Research, 27*, 329–340.
Jackson, F. (1998). *From metaphysics to ethics: A defence of conceptual analysis*. New York: Oxford University Press.
King, J. (1996). Structured propositions and sentence structure. *Journal of Philosophical Logic, 25*, 495–521.
King, J. (1998).What is a philosophical analysis? *Philosophical Studies, 90*, 155–179.
Lipton, P. (2000). Inference to the best explanation. In W. H. Newton-Smith (Ed.), *A companion to the philosophy of science*. Oxford: Blackwell Publishing.
Ludwg, K. (2007). The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy, 31*, 128–159.
Moore, G. E. (1903). The refutation of idealism. *Mind, 12*, 433–453.
Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.
Peacocke, C. (2003). Implicit conceptions, understanding, and rationality. In M. Hahn & B. Ramberg (Eds.), *Reflections and replies: Essays on the philosophy of Tyler Burge*. Cambridge: MIT Press.
Peacocke, C. (2004). *The realm of reason*. Oxford: Oxford University Press.
Plato. (1971). *Meno*. London: Penguin Books (translated by W. K. C. Guthrie).
Salmon, N. (1986). *Frege's puzzle*. Atascadero: Ridgeview Publishing.
Schroeter, L. (2004). The limits of conceptual analysis. *Pacific Philosophical Quarterly, 85*, 425–453.
Schroeter, L. (2006). Against a priori reductions. *Philosophical Quarterly, 56*, 562–586.
Soames, S. (2005). *Reference and description: The case against two-dimensionalism*. Princeton: Princeton University Press.
Sorensen, R. A. (1992). *Thought experiments*. New York: Oxford University Press.
Weatherson, B. (2003). What good are counterexamples?. *Philosophical Studies, 115*, 1–31.
Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Oxford University Press.