

Expressivism, representation, and the nature of conceptual analysis

David Plunkett

Published online: 8 July 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Philosophers often hold that the aim of conceptual analysis is to discover the representational content of a given concept such as FREE WILL, BELIEF, or LAW. In *From Metaphysics to Ethics* and other recent work, Frank Jackson has developed a theory of conceptual analysis that is one of the most advanced systematizations of this widespread idea. I argue that this influential way of characterizing conceptual analysis is too narrow. I argue that it is possible that an expressivist account could turn out to be correct as a genuine *conceptual analysis* of a genuine *concept*. I claim that since an expressivist analysis does not aim to discover the representational content of a given concept—and, indeed, might itself be based on the idea that the concept in question is not even representational in nature—the possibility of expressivist conceptual analysis shows that Jackson’s theory of conceptual analysis is incomplete as it currently stands. I conclude that Jackson needs to either shift his basic understanding of the nature of conceptual analysis or commit to a particular normative reinterpretation of his project.

Keywords Conceptual analysis · Concepts · Expressivism · Philosophical methodology · Frank Jackson · Allan Gibbard

1 Introduction

One of the standard projects that philosophers claim to engage in is the project of *conceptual analysis*. In broad terms, we might say that this is the project of better understanding the *content* or *meaning* of concepts that we use in thought and

D. Plunkett (✉)
Program in Law and Philosophy, UCLA, UCLA School of Law, Box 951476, Los Angeles,
CA 90095-1476, USA
e-mail: plunkett@post.harvard.edu

practice—e.g., FREE WILL, BELIEF, OR LAW.¹ But what exactly is involved in doing conceptual analysis? And what counts as an instance of conceptual analysis instead of an instance of another project—of, say, an instance of *substantive* argument about such things as free will, belief, and law? In *From Metaphysics to Ethics* and other recent work, Frank Jackson has attempted to answer such questions by presenting a unified so-called “neo-classical” theory of conceptual analysis (Jackson 1998). In so doing, Jackson has aimed to help make clear what conceptual analysis *is* and what roles we should (and should not) expect it to play in overall philosophical inquiry.

In this paper, I want to raise a problem for Jackson’s theory of conceptual analysis. In short, the problem is this: there are instances of conceptual analysis that Jackson’s view fails to accurately so categorize. In particular, I will argue that it seems at least *conceptually possible* that an expressivist analysis of a concept could turn out to be true. For instance, it could turn out that philosophers such as R.M. Hare, Simon Blackburn, and Allan Gibbard are correct and that expressivism is true about certain fundamental normative concepts (Hare 1981; Blackburn 1993; Gibbard 2003). Yet, on Jackson’s view, such accounts would not count as instances of conceptual analysis at all. I conclude that Jackson’s theory is thus incomplete as it currently stands and that Jackson needs to either shift his basic understanding of the nature of conceptual analysis or else commit to a particular normative reinterpretation of his project.

Before I start, let me first address two potential worries about the work that follows. The first worry is that, given how I have set things up, one might think that the central issue here is little more than a semantic debate about what falls under the labels “concept” and/or “conceptual analysis.” I think this worry is misplaced. We can ultimately choose to use words however we like: for instance, we can call certain things “concepts” and not call other things “concepts” depending on how we decide to use the term. However, the debate in this paper is not just about how we choose to use our words. In many parts of contemporary philosophy, philosophers use the language of “concepts” and “conceptual analysis” to make what they take to be important claims of a certain type—e.g., claims about conceptual content that are distinguished from *substantive* claims about the extensions of those concepts. The debate about the nature of conceptual analysis is thus more than just a debate about how best to use our terms. Rather, it is also a debate about how to understand the nature and methodology of one of the central activities that philosophers claim to engage in; and, moreover, an activity that, especially in the wake of broadly Quinean skepticism about conceptual analysis, many philosophers have thought to be in need of clarification. My concern in this paper is thus more than a semantic one—it is about the nature and methodology of the central philosophical activity of conceptual analysis.

The second worry that one might have is that, because of my focus on Jackson’s specific theory, the lessons we have to learn from my argument are limited to the particularities of one specific philosopher’s view. However, I again think this worry

¹ Following one standard convention, I will henceforth use smallcaps to designate concepts. On this convention, see Margolis and Laurence (1999).

is misplaced. Jackson's theory of conceptual analysis is representative of a neo-classical approach to conceptual analysis that has its roots in a broadly Fregean picture of concepts that has animated much of twentieth-century philosophy. As I see it, Jackson's theory of concepts is important chiefly because it is one of most sophisticated and systematic attempts to articulate an understanding of concepts that is widely accepted in large parts of philosophical practice. In some cases, this acceptance is relatively explicit—e.g., by philosophers such as David Chalmers, Phillip Pettit, David Braddon-Mitchell, and Michael Smith (Chalmers 1996; Pettit 2003; Braddon-Mitchell 2003; Smith 2004). In other cases, the theory of conceptual analysis is largely presupposed but not formally endorsed—e.g., in areas ranging from philosophy of mind to philosophy of law to formal semantics. Jackson's theory of conceptual analysis is attractive because it promises both to systematize one standard operative view of conceptual analysis and to explain its basic rationale almost solely on the basis of claims about the basic nature of language and representation that seem almost impossible to reject. However, if I am right, then Jackson's theory fails to give an adequate account of what conceptual analysis is and how it works. Insofar as his approach represents one widely assumed view of conceptual analysis, I think this in turn can help us better to understand what is missing in one dominant approach to conceptual analysis in contemporary philosophy.

2 Jackson's theory of conceptual analysis

Let's begin with Jackson's account of conceptual analysis. In *From Metaphysics to Ethics*, Jackson introduces this account by way of a general picture of the nature of language. Therefore, in order to understand Jackson's theory, I will begin with this picture.²

According to Jackson, one of the main functions of language is to *represent* things as being one way or another. Indeed, Jackson maintains, much of our linguistic activity makes sense only if we accept this as a basic thesis about the function of language. For instance, he thinks that it is what best makes sense of why, when I ask you directions to a nearby store and you tell me how to get there, I am then able actually to *find the store* by following your directions.

In rough terms, here is what Jackson thinks is going on in such a case. In responding to me with directions, you represent the world as being a certain way—namely, as a world in which the store and I are in specific locations. If I correctly grasp the information that you impart to me with this representation, I am then in a position to know where you think the store is—and, furthermore, to know where the store *actually* is if you are correct.

As with all representations, linguistic or otherwise, your statement of where I am in relation to the store effects a division between ways things might be that are in accordance with this representation and those that are not. For instance, if you tell me that the store is two blocks west of me and it turns out that this is so, then this

² In what follows, I closely follow Jackson's own way of summarizing his view in Jackson (2001).

result would be in accordance with your representation. In contrast, if it turned out that the store was actually three blocks east of me, then this would be in conflict with your representation.

Given that a chief function of words and sentences is to represent, and to represent is to make a division, one obvious question we might be interested in is what divisions certain words and sentences effect. As Jackson conceives of conceptual analysis, the aim of conceptual analysis is to answer this question. In other words, its aim is to help us better grasp what divisions our words and sentences effect when we use them to represent things as one way or another.

According to Jackson, the best way to model what information a representation provides is by employing the idea of *possible worlds*. In *From Metaphysics to Ethics*, Jackson conceives of a possible world as a *complete description of the way things might be* (considered in a so-called “two-dimensionalist” framework *either* as a counterfactual possibility or an alternative to actuality).³

With this idea of possible worlds in hand, Jackson claims that we can model any representation as a division between possible worlds that are in accordance with that representation and those that are not. In turn, this means that for a specific statement, that represents things as being a certain way, a statement such as “water is H₂O”, we can investigate *what* division among possible worlds that representation effects. This, for Jackson, is the basic project of conceptual analysis.

In most cases, Jackson claims, we are interested in this project because we are interested in *equivalent* ways of representing the world. Thus, for instance, in wondering how best to understand our term “knowledge,” we often want to know not just what division this term helps effect when used in a statement, but, moreover, how we might make the *same* division among possible worlds using other terms instead. For instance, we might wonder whether or not the statement “X is justified true belief” makes the same division as the statement “X is knowledge” or whether, as the Gettier cases suggest, the two statements effect different divisions. To offer a specific conceptual analysis, then, often is to propose that two different representations actually effect the same division among possible worlds; or, in other words, it is to propose that both representations pick out the same possible pattern in nature.

In presenting this picture of conceptual analysis, Jackson usually presents it in terms of *words*. However, Jackson’s theory of conceptual analysis is ultimately more than just a theory about words. This is because it is also a theory that can apply to other things we use to represent—for example, maps. Furthermore, given that we *use* words to express our thoughts, and, moreover, that we can use the *same* word to express different thoughts if we so choose, it is also a theory about the *thoughts* we use words or other representational items to express.⁴ More specifically, if we think

³ It should here be noted that this description of possible worlds gets more complicated insofar as Jackson endorses the idea of *centered worlds* that include *de se* information. However, this detail is irrelevant for the purposes of this paper.

⁴ I am here assuming a certain theory of how to individuate words: namely, one on which it is possible for the same word to express two different concepts. However, as many contemporary linguists seem to, one might prefer a more fine-grained individuation of words according to which this is not so. That is: one might prefer a theory according to which if a word *means* something different from another word, then it

of token concepts as the mental side of token words, Jackson's theory is ultimately a theory about how to understand *concepts*.⁵

According to Jackson, what it is to offer a conceptual analysis of a given predicate-type concept such as WATER or KNOWLEDGE is to offer an account of the particular divisions among possible worlds that are effected by statements using that concept.⁶ This can be fleshed out as follows. In basic terms, for Jackson, what it is for a subject to possess a given concept C is for that subject to possess a *conditional ability* to identify the extension of that concept C in a given possible world, given sufficient information about that possible world and sufficient reasoning (Chalmers and Jackson 2001). For instance, take the concept KNOWLEDGE. According to Jackson, possession of the concept KNOWLEDGE gives one the capacity to evaluate conditional statements of the form "PW \supset K", where "PW" is a statement that gives sufficient information about a possible world and "K" is a statement that both uses the concept KNOWLEDGE and characterizes its extension (such as the statement "justified true belief is knowledge"). We can thus say that a concept C determines a function from possible worlds to extensions. Or, more accurately, given Jackson's two-dimensionalism, we can say that a concept C determines a function from possible worlds to extensions *given* sufficient information about the possible world that we are supposing is actual. Furthermore, for the purposes of this paper, we can say that, for Jackson, the *content* of a given token concept is the particular function from possible worlds to extensions that it determines.⁷ In turn, holding fixed the idea we want to individuate concepts by the specific division among possible worlds that a given concept helps to effect when employed in statements, we can say that, for Jackson, a given token concept is individuated by the *exact* function from possible worlds to extensions that it determines.

To sum up, we can thus say that Jackson presents us with the following picture: a chief function of language is to *represent* the world as one way or another by expressing thoughts that we have about how the world is, and *concepts* are representational components of our thoughts that help us to represent the world as one way or another. In turn, given the basic possible worlds framework that Jackson uses to understand representation, we can say that the *content* of any given concept

Footnote 4 continued

should be thought of as a different entry in a speaker's lexicon and hence should be thought of as a different word. For instance, on such a theory of word individuation, the word "bank" we use to refer to the place where we deposit money would be a different word than the word "bank" we use to refer to the side of a river. For my purposes here, this choice about how to individuate words will not matter. This is because, if we think of words as individuated in such a fine-grained way, we can still think of doing analysis on the concepts that these words express even if we cannot use the same word to express different concepts.

⁵ This point is underscored in David Braddon-Mitchell's way of developing Jackson's basic understanding of conceptual analysis in (Braddon-Mitchell 2003).

⁶ I use the term "predicate-type" concept here to signal that Jackson's theory is aimed at concepts such as those expressed by the terms "dog" and "chair" and not at such things as logical or modal operators.

⁷ It should be noted that the term "content" is used in *many* different ways in the philosophy of language. In using the term "content" in this way, I do not intend to take a stand on broader debates about how to use the term in general. Rather, I am only doing so in order to provide a helpful way of understanding Jackson's theory.

is a function from possible worlds to extensions, and that the project of *conceptual analysis* is the project of figuring out what functions from possible worlds to extensions are determined by different token concepts.

3 Gibbard's expressivist analysis of normative concepts

With Jackson's understanding of concepts in hand, I now want to introduce what I think is a serious problem for his theory: namely, that it rules out the possibility of an *expressivist* analysis of a concept and, hence, I argue, that it fails to capture everything that we currently count as a genuine instance of conceptual analysis. Thus, insofar as Jackson's view of conceptual analysis is meant to capture and to elucidate what we currently think of as the philosophical project of conceptual analysis, it is in fact an inadequate picture of this project.

In order to see this problem, we need first to appreciate what an expressivist analysis of a concept is and/or at least might purport to be. In order to effectively do so, I will here consider one of the most advanced and influential contemporary expressivist accounts—namely, Allan Gibbard's expressivist account of normative concepts (i.e., concepts about what *should* be the case).

According to Gibbard, normative concepts about what should be the case—including, for him, concepts such as MORALLY RIGHT, RATIONALITY, and REASON FOR BELIEF—are irreducibly *practical* concepts that, as he puts it in *Thinking How To Live*, concern *planning* questions of what to do, think, or feel. According to him, the basic role of these planning concepts is *not* correctly to represent a matter of fact either truly or falsely. Rather, Gibbard argues, the basic role of these planning concepts is to engage planning-like mental states. As he puts it in *Reconciling Our Aims*: “ought thoughts are like plans” (Gibbard 2008, p. 19). For instance, in rough terms, Gibbard thinks that when I think the thought “I should donate more money to Oxfam if I become richer,” I am in a planning-like mental state about what to do if I become richer—namely, to donate more money to Oxfam.

In turn, on Gibbard's expressivist theory, when I utter a sentence such as “I should donate more money to Oxfam” I *express* this mental state. This is, for Gibbard, in contrast to *stating* that I am in that mental state. To see how this distinction is supposed to work, consider my yelling “Hoorah Red Sox!” in a Red Sox game when they are winning. In such a case, I yell this because I have a pro-attitude towards the Red Sox winning. However, I do not yell this in order to *state* that I have this attitude and to get you to believe something about my mental life—as it would be, for instance, if I uttered the sentence “my favorite baseball team is the Red Sox” in a normal context. Rather, I yell this in order to *express* my pro-attitude toward the Red Sox. Just as I express this pro-attitude in this case, so too does Gibbard think that I express my planning-like mental state when I make a normative claim.

For Gibbard, the planning-like mental states involved in normative thought and talk are *noncognitive* in the sense that rather than aiming to fit the way the world is by producing a correct representation, they aim to change the world to make it fit their representation of how it should be. That is: they have a so-called “world-to-mind”

direction of fit like desires rather than a so-called “mind-to-world” direction of fit like beliefs. In basic terms, this is what we can say makes his position *expressivist* rather than *descriptivist*.

At the same time, however, Gibbard argues in *Thinking How to Live* that the planning-like mental states involved in normative thought exhibit a structure that mimics the structure of descriptive belief. More specifically, Gibbard thinks that given the temporal nature of plans, they are subject to constraints of coherence and consistency in a way that is similar to beliefs. Because of this, he thinks that we can meaningfully talk, at least in a so-called “quasi-realist” way, of normative thoughts as consisting in *belief* with representational content that can then be true or false according to a minimalist account of truth-aptness; and, thus, he thinks, there is a sense in which we can also think of his view as a “cognitivist” one.⁸ This is why, for Gibbard, despite the fact that plans have a different direction of fit than prosaic descriptive judgments, we can think of these plan-like states as *judgments* despite the fact that we might normally think of judgments as, by definition, having content about the way things *are*. However, what is important to stress is that, on Gibbard’s picture, we earn the right to talk this way about the genuine truth and falsity of normative beliefs *because* of features of the noncognitive mental state involved in normative thought, rather than the other way around. In short, the noncognitive attitude is explanatorily primary.

In order to appreciate how Gibbard’s quasi-realist view thus differs from a descriptivist alternative, consider that, on Gibbard’s theory, there is always an equivalent descriptive way of picking out the property of being *the thing to do* that is picked out in an expressivist judgment. For instance, holding fixed a substantive normative position that agents should maximize overall utility, we could spell out a naturalistically specifiable predicate that picks out all and only those possible worlds in which agents act to maximize utility. We could then use this predicate to form straightforward descriptive beliefs about whether or not agents act how they should. However, for Gibbard, the important point is that if we were to pick out the very same property in this other way, we would *not* be doing so using genuinely *normative* concepts. Indeed, on his view, in contrast to Jackson’s, one can show that two concepts have the same extensions across all possible worlds *without* thereby showing that these two concepts are conceptually equivalent. This is because Gibbard holds that even if there is an important “quasi-realist” sense in which we earn the right to talk of normative beliefs, normative concepts are different *in kind* from descriptive concepts: they are defined not by having representational content of one sort or another, but rather by the planning-like noncognitive attitudes we engage when we use these concepts. This is why when Gibbard discusses a normative concept such as OUGHT TO BE DONE, he does *not* tell us about the specific division across possible worlds that this concept helps to effect—let alone offer an analysis of this concept in terms of an equivalent way of effecting the same division. Rather, he gives us what he calls an “oblique” analysis that explains this and other normative concepts in terms of the planning-like noncognitive mental state that one is in when one thinks that something ought to be done. This “oblique” analysis does

⁸ See Blackburn (1993) on the idea of “quasi-realism.”

not tell us anything like the truth-conditions for statements that employ this concept. Rather, it gives us an account of what we *do* with these concepts in practice. Following this, given the expressivist account that Gibbard gives of what we in fact do with these concepts, he thereby gives us what we might call “felicity-conditions” for using these concepts in a proper way.

To sum up, we can thus say that Gibbard presents us with the following picture: normative *concepts* are non-representational concepts that one uses when engaged in the planning-like activity that is distinctive of normative thought and talk. In turn, the *content* of any given normative concept *cannot* be understood as a function from possible worlds to extensions but rather must be understood obliquely in terms of the planning-like mental state one is in when one engages in normative thought and talk.⁹

4 Why expressivism is a problem for Jackson’s theory of conceptual analysis

In order to appreciate why Gibbard’s expressivism poses a problem for Jackson’s theory of conceptual analysis, let’s start with the following question: how should we think of the status of Gibbard’s expressivist account of normative concepts? More specifically, should we think of Gibbard as offering a *conceptual analysis* of normative concepts?

As he himself suggests, I think there is good reason to hold that this is an accurate account of what he is doing. This is because, rather than trying to show us what things actually *are* normative for us, what Gibbard is trying to show is *what’s at issue* when we engage in normative thought and talk. Or, to put it another way, it looks as if he is trying to understand the *nature* and *meaning* of normative concepts as opposed to offering substantive views *using* those concepts.

We might thus paint the picture of what is going on here as follows. If we think of Gibbard as doing conceptual analysis, then we should say that there are *two* distinct questions that one needs to ask when engaged in the analysis of a particular token concept. In broad terms, these are (1) for what end do we use this concept in practice? and (2) given the end for which we use this concept in practice, what further can we say about the content of this concept that explains how it does this job? In thinking of all concepts as representational in nature, Jackson takes the first question as already answered in advance for all concepts: and, thus, he is able to think of conceptual analysis as only involving the explanation of how the content of a given concept allows us to make statements using that concept that effect a certain division among possible worlds. According to Gibbard, however, this is a mistake. In short, we might say, Gibbard thinks that neo-classical theories of concepts like Jackson’s have had an overly narrow picture of what sorts of functional roles concepts can play. Once we see the other roles concepts can play, Gibbard can say,

⁹ It should be noted that, on Gibbard’s account, for any *thick* normative concept that involves both normative and descriptive criteria for correct application, this concept will have some representational content. This is important to keep in mind when understanding how an expressivist view such as Gibbard’s is meant to work.

we can see why the type of answers we give to the second question we need to ask in conceptual analysis can vary widely depending on what *type* of concept is involved. This is why, Gibbard can maintain, his expressivist theory counts as an instance of conceptual analysis despite looking so different in kind from the paradigmatic conceptual analyses that Jackson considers in formulating his theory.

However, according to Jackson's account of conceptual analysis, what Gibbard is doing is *not* conceptual analysis. In short, this is because, on Jackson's view, there simply is no room for something to be a genuine concept even susceptible to analysis if it is not the sort of thing that can be identified by the function from possible worlds to extensions that it determines.

This, though, is a problem for Jackson's theory of conceptual analysis. For even if Gibbard's expressivist theory is wrong, it seems at least *conceptually coherent* to imagine that he is correct. This means that, whatever our concept CONCEPT is, we can at least imagine that some concepts are expressivist in the way that Gibbard proposes—and, moreover, that would then be correctly analyzed along the lines that Gibbard suggests. Gibbard's expressivist analysis thus poses a problem for Jackson's theory of conceptual analysis. For, at least as he presents it in *From Metaphysics to Ethics*, Jackson is not simply aiming to describe a new activity that he has come up with and decided to call "conceptual analysis." Instead, it seems, his aim is to capture and to elucidate the central projects of conceptual analysis that philosophers are *already* engaged in; projects that they either present explicitly under the label "conceptual analysis" or else projects that, as Jackson puts it, we can recognize as projects of "closet conceptual analysis" once we have the paradigmatic instances of conceptual analysis in sight (Jackson 1998, p. vii). Therefore, insofar as Jackson's project has this descriptive aim, and we count Gibbard as doing conceptual analysis, it appears that Jackson's account of conceptual analysis is incomplete.

Indeed, I think, we can see this problem for Jackson's theory even *by its own lights*; or, more specifically, we can see this as a problem for his theory insofar as it purports to be a good account of what we are already thinking of when we think of something *as* an instance of conceptual analysis. According to Jackson, a successful conceptual analysis of a given concept should identify the possible pattern in nature that this concept covers. Yet, if our concept of CONCEPTUAL ANALYSIS *does* include what Gibbard is doing in its extension, whereas Jackson's account of conceptual analysis does not, then this means that Jackson fails to give us an accurate account of the content of the concept CONCEPTUAL ANALYSIS.

To see the import of this point, consider one of the things that is taken for granted by many philosophers currently working on expressivism: namely, that if there are *any* expressivist concepts in any world, then surely we should have *some* of them in the actual world given how interesting and seemingly useful they would be.¹⁰ Call this the *instantiation premise*. If true, this premise has important implications. Most importantly, suppose that Jackson understands his theory of conceptual analysis not in the first instance as an account of the concepts CONCEPT and CONCEPTUAL ANALYSIS, but rather as a substantive theory of what conceptual analysis consists of in *our*

¹⁰ This assumption is reflected in Gibbard's emphasis in Gibbard (2003) on establishing the *possibility* that expressivism could be true for a given area of thought and talk.

actual world given the nature of actual concepts in our world, or, more broadly, given certain constraints on how an actual language in a world such as ours needs to work. If the instantiation premise is correct, then even if Jackson were to understand his theory in this way, then the success of Jackson's theory as a full systematic account of conceptual analysis would still depend on the issue of whether or not expressivist conceptual analysis is *conceptually possible*.

5 The conceptual coherence of expressivism

In this section, I will now consider what I think is the most promising response to my argument on Jackson's behalf. In broad terms, the response is to argue that, given the nature of concepts, it is not in fact conceptually coherent to imagine that Gibbard's expressivist analysis could turn out to be true.

How might such an argument best be motivated? Consider again the statement of "hoorah Red Sox!" that I mentioned earlier when introducing the basic form of expressivism. When a speaker utters this phrase, and thereby expresses her pro-attitude to the Red Sox, do we think that she is using a *concept* HOORAH? Even if we think that "hoorah" *means* something and that there are conditions for using "hoorah" properly, it seems plausible to hold that we do not think of it as a concept—and, moreover, to hold this on the basis of the fact that "hoorah" has no representational content. If so, we might claim with Jackson that the only things that can be genuine concepts are those things with representational content.

One good reason to endorse this line of thought is that even if the statement of "hoorah Red Sox!" can start a type of *conflict* with another person—e.g., the conflict that might ensue when a Yankees fan yells back "I hate the Red Sox!"—it does not look as though it makes possible what we think of as a *canonical disagreement* of the same sort that might emerge from straightforward factual statements such as "there are five chairs in this room" or " $2 + 2 = 4$." Part of the job-description of concepts, we might say, is to explain how these two types of conflict differ from each other; and, moreover, how they differ from each other *even if* someone like Gibbard is right that they both can be called "disagreement" in one sense of the word in our natural language. Expanding on this point, we might say that although language functions in a variety of ways, we use the idea of *concepts* to explain how it works in the important and distinguished way it does when we engage in the sort of straightforwardly truth-conditional representational activity that Jackson is concerned with. For, by understanding concepts in roughly the way that Jackson thinks of them, we are able to show how there is something *in common* between the thoughts of agents engaged in canonical disagreement that is lacking when agents simply have divergent plans or noncognitive attitudes.

I take this to be a central line of thought that supports Jackson's broadly classical or Fregean notion of concepts; and, moreover, one that can reasonably be thought of as part of the reason for this type of neo-classical approach in the first place. For, by thinking of restricting concepts to those things that can figure in the truth-conditions of propositions in a traditional sense, we can use them as a tool for explaining the

type of representational activity that figures in what I characterized as “canonical disagreement.”

Indeed, Jackson might hold that, to think otherwise about concepts, Gibbard owes us at least the *outline* of his account of the concept CONCEPT such that non-representational entities *could* be concepts. Without it, Jackson might argue, it is better to think that what he has given is the correct analysis of the concept CONCEPT and that Gibbard is considering things that are somehow different from concepts—perhaps things that, like logical or modal operators also contribute to the *meaning* of statements but that do so in a fundamentally different way than concepts. Furthermore, perhaps these other things function in much the same way as concepts on the surface—and hence might earn the right to be thought of as “quasi-concepts” as part of Gibbard’s “quasi-realism.” Yet, Jackson might maintain, given that Gibbard’s *explanation* of how these quasi-concepts serve this function points to a radical difference between them and standard concepts such as DOG or CHAIR, such “quasi-concepts” should still not be thought of as genuine *concepts*. To push this line of thought even further, Jackson can then try to save his own theory of conceptual analysis as a theory of *conceptual analysis* proper and claim that, at best, what Gibbard is doing is conceptual* analysis of things that are concept*s. To put it another way, Jackson can hold that Gibbard might be making a genuine contribution to the theory of *meaning* of terms like “ought” and “should,” but that this does not warrant him to think of him as engaged in genuine conceptual analysis.

For the reasons I outlined above, I think that this response is philosophically well motivated. However, it faces a serious problem. The problem is this. In order for this line to work, Jackson would need to convince us that what look to us from our current vantage point to be genuine normative *concepts* like OUGHT and SHOULD would turn out *not* to be concepts if we discovered that Gibbard’s expressivism were true. Or, to put it another way, on Jackson’s account, the only way that we would have genuine concepts of OUGHT and SHOULD would be if some form of metanormative *descriptivism* turned out to be true according to which such concepts are genuinely representational in nature. However, I think that our belief that these are genuine concepts would hold *even if* we discovered that descriptivism turned out to be wrong. What we would conclude, I think, is that we had the wrong analyses of these concepts—not that we lacked such concepts altogether. In some cases, then, Jackson might be right that when we discover the broadly “noncognitive” nature of a domain of our thought and talk, we will conclude that it is not trading in genuine concepts. However, I don’t think he should conclude that this means that *anytime* one shows an area of thought or talk to be non-cognitivist one thereby shows that the purported concepts in question are really concept*s. This does not follow. For instance, it might be that those domains of non-cognitivist thought and talk that earn the right to be “quasi-realist” involve genuine concepts whereas those that are not “quasi-realist” do not. In turn, this would help explain why we have a concept of OUGHT but not of HOORAH. Putting this specific conjecture to the side, the basic point is this: I think that our intuitions about the genuine *conceptual* nature of normative concepts (even in the face of the potential truth of Gibbardian expressivism) are strong enough to resist this proposed solution on Jackson’s behalf.

Furthermore, another problem with this proposed solution is that it might solve the problem in name only. For even if one grants that what Gibbard is doing is conceptual* analysis of concept*s, rather than conceptual analysis of genuine concepts, this might mean that the real problem then just becomes how to identify what conceptual* analysis has in common with conceptual analysis. For if there is *some* concept, call it CONCEPT+, whose extension covers both expressivist concept*s and descriptivist concepts, it might be that this is really the concept that has confused us—and not the concept CONCEPT. Similarly, suppose that one rejected the tight connection that Jackson assumes between the concepts CONCEPTUAL ANALYSIS and CONCEPT such that a philosopher such as Gibbard could be engaged in genuine conceptual analysis despite there not being any expressivist *concepts* there at all. Call this concept CONCEPTUAL ANALYSIS +. If so, then there is still a question of what practitioners of CONCEPTUAL ANALYSIS + are doing and what they are studying if not concepts. On the plausible assumption that *this* broader activity is the thing that initially confused us—it was the activity, in other words, that we were hoping for Jackson’s work to illuminate—we would then have to conclude that his work explains only one limited part of it. In other words, although we initially put the problem *in terms* of concepts and conceptual analysis, this wording might not be correct—and, yet, call it what you want, the same basic philosophical problem remains.

This does not mean that there is nothing more to be said on behalf of this basic response. For there might be further arguments to give on why the idea of a genuine expressivist conceptual analysis of a genuine concept is in fact conceptually incoherent. For instance, consider the fact that expressivists have traditionally had a problem with distinguishing *negation* from *rejecting believing* (Unwin 2001). In recent years, a number of proposals to this solution have been proposed that look like quite promising—for instance, those by Schroeder (2008b) and Dreier (2006). Suppose, however, that these solutions fail and that there is in fact no way for an expressivist to adequately explain the possibility of negation. If so, then one might maintain that expressivism not only fails to adequately capture disagreement but is in fact incoherent because it fails even to describe a possible language.¹¹ In such a case, we might then conclude that an expressivist conceptual analysis is ruled out as conceptually incoherent from the start. To put it another way, we might say that, to be a concept is to be capable of being subject to negation. Thus, if an expressivist concept is not in fact subject to negation, then it is in fact no concept at all—and thus cannot be something that is the subject of genuine conceptual analysis.

If such a response were to work, it would thereby show that there is *not* a possible pattern in nature that could be picked out by a concept that contains both

¹¹ There are, of course, other arguments besides these that one might use to establish the claim that expressivism fails to describe even a possible language. Jackson himself has advanced such arguments. For instance, in “A Problem for Expressivism,” Jackson and Pettit argue that a so-called “Lockean” picture of how ethical terms such as “good” acquire their meaning undermines the coherence of expressivism (Jackson and Pettit 1998). For critical discussion of this argument see Smith and Stoljar (2003) and Jackson and Pettit (2003). For discussion of other problems in making coherent the idea of an expressivist language, see Schroeder’s recent *Being For* (Schroeder 2008a).

expressivist and descriptivist concepts—for, in short, if expressivism is not only wrong but impossible, then there is no possible pattern in nature that would include both descriptivist and expressivist concepts. As we know from the history of attempted conceptual analyses of such simpler concepts as WATER, there are almost always unobvious parts of a given concept's content. In essence, if Jackson were to go this route, he would basically be saying that there are also unobvious parts of our concepts CONCEPT and CONCEPTUAL ANALYSIS that we see from doing more high-level work in metaethics and the philosophy of language.

Alternatively, once one has in hand arguments against the coherence of expressivism such as the negation problem, Jackson could put forward these arguments against the coherence of expressivism as ones not about the *conceptual* coherence of expressivism, but rather as ones about the viability of expressivism as an account of concepts that we employ in the actual world. In other words, he could reject the *instantiation premise* I introduced earlier in this paper and then claim that he does *not* aim to give an account of all conceptually possible forms of conceptual analysis, but rather of the form that would work given the sorts of concepts we employ in the actual world.¹² In essence, if Jackson were to go this route, he would, as in the previous response I just considered, again be drawing on more high-level work in metaethics and the philosophy of language to make a claim about the viability of expressivist conceptual analysis that would not be immediately obvious based on initial reflection about our concepts CONCEPT and CONCEPTUAL ANALYSIS. In this case, the claim would be non-obvious not because of anything having to do with the content of these concepts, but rather because of the substantive facts about how thought and language work in the actual world.

Are such arguments against expressivism strong enough that we should go one of these two routes? Perhaps we will ultimately conclude that they are. But I do not think we are currently in a position where we should conclude that. The first reason why is that, although the general forms of these two strategies might be viable, it looks as though there might in fact be strong responses in the offing on behalf of expressivism. These include promising proposed solutions to expressivism's negation problem—e.g., those offered by Schroeder and Dreier. It is thus not at all clear that either of these proposed strategies could be made to work.

This points to a deeper reason to be skeptical of both these proposed ways of defending Jackson's theory of conceptual analysis. If what Jackson is doing is trying to give us a broad characterization of the nature and methodology of conceptual analysis, then, in formulating a theory that explains what unifies the type *conceptual analysis*, he should set pretty high evidential standards for ruling out the possibility that a leading token of purported conceptual analysis might be correct (and hence taking it as one of the data points his theory needs to account for). Any plausible set of standards here will, of course, need to include the assessment of the viability of that given token conceptual analysis as well as, in the case where there are experts working on the area of that given conceptual analysis, an assessment of the state of play in that area in order to gain indirect evidence about how best to evaluate the epistemic status of that analysis. Given that expressivism is, by almost all lights of

¹² See Schroeder (2008a) for one of the strongest recent arguments of this type.

those engaged in contemporary metaethics, currently one of the “going options” on offer, I think we are not in a position to rest the viability of our theory of conceptual analysis on the highly contentious claim that expressivism is either *impossible* or at least not at all viable as an account of *any* area of natural language in the actual world.

6 Conclusion

If what I have argued so far is correct, then I think we should conclude that Jackson’s theory of conceptual analysis is inadequate in virtue of the fact that we think a Gibbard-style account of genuinely expressivist concepts should count as *some* form of conceptual analysis. In essence, this means that Jackson’s current account of conceptual analysis is an *unfinished* one at best. This, of course, would not undercut the idea that Jackson has nonetheless accomplished something important in the work he has done so far. For instance, Jackson could claim that he has captured how to understand an important *species* of concepts and, in turn, captured what is going on in an important part of conceptual analysis. Nonetheless, even if this were so, it still leaves unanswered the questions we started with about the nature and methodology of conceptual analysis as a whole.

Furthermore, if one accepted this basic line of thought, one can see how attempting to make further progress on this front is going to be difficult. For instance, suppose Jackson wants to keep his basic idea that concepts as a *type* are identified as things that play a certain functional role in our thought. If descriptive concepts are defined by playing a certain functional role in our thought—namely, a representational one—but they turn out to be only a *species* of concepts, what sort of functional role defines the *genus* of concepts in general? Given how distinct the roles of concepts will turn out to be if expressivist concepts are possible, it appears that it will be difficult to find a satisfactory answer to this question—and, moreover, one that will necessarily lead us to an overall unified theory of conceptual analysis that departs significantly from Jackson’s current one.

One reconciling option here worth mentioning would be for Jackson to draw on Gibbard’s work and adopt a broader understanding of “representation” than he currently does. For instance, if one thought it were possible to represent either over possible worlds or over Gibbard-like plans, then Jackson could still think of *all* concepts as fundamentally in the functional business of possibility-carving: it would just turn out that they are helping to carve possibilities over different things. In turn, Jackson could then still see the project of conceptual analysis as discovering the specific possibility-carving profiles of different concepts: again, it would just be that there are different ways of having a possibility-carving profile depending on what sorts of things one is carving a possibility over.¹³ I leave it open for future work to determine whether or not such a proposal is in fact plausible. For now, I mention this proposal to signal the type of revision that I think might be successful—and,

¹³ Thanks to Andy Egan and Scott Shapiro for originally raising this line of thought to me in separate conversations.

moreover, to underscore that such a revision would indeed constitute a significant change to Jackson's current understanding of conceptual analysis.

In the end, perhaps there are no simple answers as to what unifies all concepts *as* concepts and to what unifies all conceptual analyses *as* conceptual analyses. However, by the lights of Jackson's own theory, insofar as we are competent users of a given token concept (e.g., CONCEPT), then we know that there must be *some* possible pattern in nature that we are picking up on and that is implicitly guiding our application of the concept. If I am right, then, the upshot of my argument in this paper is this: even if this pattern is quite complicated and heterogeneous in certain respects, and hence hard to correctly identify, we can still conclude that Jackson himself has not yet correctly captured it with his theory of conceptual analysis. In other words, we can still conclude that his theory is incomplete and, moreover, would require significant fundamental revision to fulfill its basic explanatory aim as I have understood it in this paper.

That being said, there is nonetheless another option for Jackson that I want to mention in conclusion. In this paper, I have treated Jackson's theory of conceptual analysis as a descriptive one that aims to correctly identify and to elucidate the existing practice of conceptual analysis in philosophy. However, Jackson could deny this by giving a decidedly *normative* reading of his project. In other words, he could claim that in giving a theory of conceptual analysis, he is primarily concerned not with the project that philosophers currently *are* doing under the label "conceptual analysis," but rather with a project that they *should* be doing—and, moreover, one that bares enough resemblance to current projects of so-called "conceptual analysis" that he can use the label in a productive way.

There are some good grounds for taking this view of Jackson's project seriously. If it turns out that Gibbard's "quasi-realist" expressivism is correct about normative concepts, then, in essence, what we would discover that such concepts behave in almost all respects like descriptive concepts at the surface level but in fact are *massively* different from normal descriptive concepts. When we discover such a basic bifurcation in the seemingly unified group of things we have called "concepts," we might very well come to think that our current concept CONCEPT picks out two *very* different sorts of entities rather than picking out something akin to a natural kind. If so, as with many of our folk concepts that we discover fail to carve nature at its joints, we might then come to think that we should not worry too much about what our current concept CONCEPT is, or with any concept that picks out *both* descriptive and expressivist concepts, but rather with independently better understanding the two different kinds of entities that we have discovered—and then, in turn, developing the proper methodologies to study each of them.¹⁴

With this sort of idea in hand, Jackson could argue that he has given us the methodology and the metaphysical rationale for one very sensible project for philosophers to pursue—and, moreover, one that captures what is going on in many of the paradigmatic instances of existing so-called "conceptual analysis." The fact that there might be another sort of project that would be appropriate for studying

¹⁴ Thanks to Tim Sundell and Kenny Easwaran for originally raising this line of thought to me in separate conversations.

another very different sort of entity, namely the things that Gibbard thinks of as “expressivist concepts,” is thus largely beside the point. For, in short, in proposing this methodology, Jackson would be giving a *normative* argument for one sort of project to pursue, one that he calls “conceptual analysis.” This is consistent with thinking that there might be other, related projects that one should pursue in philosophy that concern things other than what Jackson thinks of as “concepts.”

If Jackson were to pursue *this* way of reading his project then he would have what I think is a successful defense against my argument in this paper. For, in effect, he would thereby show that my argument essentially misses its target altogether given the normative aim of the project. However, it is not clear that Jackson *should* endorse this re-interpretation of his work. This is because, as is implicitly suggested by the very choice of the subtitle “A *Defense* of Conceptual Analysis” for his book *From Metaphysics to Ethics*, part of what makes Jackson’s theory initially attractive seems to be its promise of giving a descriptive account that unifies and defends our *existing* projects of conceptual analysis. This means that, if I am correct in this paper, Jackson ultimately turns out to face a dilemma: namely, either he must admit that his theory of conceptual analysis is incomplete, in which case it will require significant revision to finish, or else he must commit to a specific normative reading of his work that might turn out to undermine the appeal of his theory. Moreover, insofar as Jackson’s theory of conceptual analysis represents a broader neo-classical approach to conceptual analysis that is often taken for granted in contemporary philosophy, this dilemma also creates a further dilemma for philosophers relying on this style of approach: namely, either they must admit that they have been relying on too narrow a conception of their philosophical task, in which case they must then become more attentive to expressivism as a form of genuine conceptual analysis, or else they must understand their own conceptual work to ultimately rest on a sort of normative argument that might in fact undermine the philosophical appeal of their proposed conceptual analyses.

Acknowledgments My work in this paper has benefited enormously from a wide range of comments and conversations. I would particularly like to thank Frank Jackson, Allan Gibbard, Peter Railton, David Braddon-Mitchell, Andy Egan, Tim Sundell, Kenny Easwaran, Scott Shapiro, Eduardo Garcia-Ramirez, Nadeem Hussain, Sam Liao, Daniel Nolan, Tristram McPherson, Stephen Darwall, Howard Nye, John Ku, Michael Smith, Tamar Schapiro, Peter Galison, Dave Chalmers, Sharon Street, Daniel Stoljar, Sarah Buss, Adam Plunkett, Richard Holton, Ben McKean, Alex Gourevitch, Alejandro Perez Carballo, Paolo Santorio, Christine Korsgaard, Boris Kment, Arnold Davidson, and an anonymous referee at *Philosophical Studies*.

Open Access This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Blackburn, S. (1993). *Essays in quasi-realism*. Oxford: Oxford University Press.
 Braddon-Mitchell, D. (2003). Qualia and analytical conditionals. *Journal of Philosophy*, 100(3), 111–135.
 Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.

- Chalmers, D. J., & Jackson, F. (2001). Conceptual analysis and reductive explanation. *Philosophical Review*, 110, 315–361.
- Dreier, J. (2006). Negation for expressivists: A collection of problems with a suggestion for their solution. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics*. New York: Oxford.
- Gibbard, A. (2003). *Thinking how to live*. Cambridge, MA: Harvard University Press.
- Gibbard, A. (2008). *Reconciling our aims: In search of bases for ethics*. Oxford: Oxford University Press.
- Hare, R. M. (1981). *Moral thinking*. Oxford: Oxford University Press.
- Jackson, F. (1998). *From metaphysics to ethics: A defence of conceptual analysis*. Oxford: Clarendon.
- Jackson, F. (2001). Precis of from metaphysics to ethics. *Philosophy and Phenomenological Research*, 62(3), 617–624.
- Jackson, F., & Pettit, P. (1998). A problem for expressivism. *Analysis*, 58(4), 239–251.
- Jackson, F., & Pettit, P. (2003). Locke, expressivism, conditionals. *Analysis*, 63(1), 86–92.
- Margolis, E., & Laurence, S. (1999). *Concepts: Core readings*. Cambridge, MA: MIT Press.
- Pettit, P. (2003). *Rules, reasons, and norms: Selected essays*. Oxford: Oxford University Press.
- Schroeder, M. (2008a). *Being for: Evaluating the semantic program of expressivism*. Oxford: Oxford University Press.
- Schroeder, M. (2008b). How expressivists can and should solve their problem with negation. *Nous*, 42(4), 573–599.
- Smith, M. (2004). *Ethics and the a priori: Selected essays on moral psychology and meta-ethics*. Cambridge: Cambridge University Press.
- Smith, M., & Stoljar, D. (2003). Is there a lockean argument against expressivism? *Analysis*, 63(1), 76–86.
- Unwin, N. (2001). Norms and negation: A problem for Gibbard's logic. *Philosophical Quarterly*, 51(202), 60–75.