

## Platitudes and Metaphysics

Forthcoming in Braddon-Mitchell, D, and Nola, R. (eds) *Naturalism and Analysis*. MIT Press, Cambridge MA.

Penultimate Draft 25/11/05

Abstract: One increasingly popular technique in philosophy might be called the "platitudes analysis": a set of widely accepted claims about a given subject matter are collected, adjustments are made to the body of claims, and this is taken to specify a "role" for the phenomenon in question. (Perhaps the best-known example is analytic functionalism about mental states, where platitudes about belief, desire, intention etc. are together taken to give us a "role" for states to fill if they are to count as mental states.) We then look to our best theory of the world to see where this role is satisfied, if at all. Unfortunately, the platitudes analysis, so characterised, does not seem to help when we are doing fundamental metaphysics—when we want to know what, at base, our world is like (and not merely where things like e.g. the mental would be found in an already-specified ontology). Nevertheless, I will argue that the platitudes analysis, properly understood, does have the materials to help us answer questions in fundamental metaphysics as well. I will explore three different ways it can do so.

Whatever philosophers say they are doing, one thing that many philosophers seem to be doing is "analysis" from the armchair. They put obvious truths about a subject matter together and see what results, or they engage in thought experiments to try to make their view plausible or a rival implausible; they draw out deductive (or near-deductive) consequences of their own view or their opponent's view, and take themselves to have established something important: all in all, much of this does not seem to involve laboratories or surveys.

Furthermore, these processes go on not just when philosophers are trying to work out the meanings of expressions, or the applicability of concepts, but also when they are trying to solve problems that are apparently about the world, rather than our descriptions of the world. Philosophers interested in the nature of rational desire, or action, or weakness of the will; or causation, or properties, or laws of nature; or epistemic justification or knowledge, will sometimes engage in activity that looks more like analysis of concepts than empirical psychology, experimental physics, or empirical sociology.

A group of theories collectively (and originally derisively) labeled the "Canberra Plan" have offered a story about what is going on here.<sup>1</sup> According to this story, what is going

---

<sup>1</sup> The expression is originally due to Price and O'Leary Hawthorne 1996. A more recent appearance of

on is indeed conceptual analysis, though of a slightly different sort to some traditional conceptions of that activity, and the conceptual analysis helps answer questions apparently about things rather than about the meanings of particular words or the application of particular concepts by solving “location problems”.<sup>2</sup>

In outline, the process is as follows. We select something we would like a philosophical analysis of: causation, colour, free will, beliefs, moral value, or whatever. Then we collect together the “platitudes” concerning our subject matter. I am not sure everyone means exactly the same thing by this technical term (see below), but as a first pass platitudes are claims about the subject matter that reflect our ordinary use of the term (or specialists’ ordinary use of the term, if we are after a philosophical analysis of a term of art, like “gene” or “electron” or “social class”). Once we have these platitudes assembled, we conjoin them, and use the resulting information to define a “theoretical role” for the thing we are interested in. Colours are the properties of surfaces of objects that cause certain kinds of experiences through producing changes in our eyes in suitable illumination, etc., or whatever it might be. It is often useful to represent this theoretical role with a “Ramsey sentence”: an existentially quantified sentence capturing all, or most, of the information in the platitudes. (The Ramsey sentence for the property of being a free action may start, e.g. with “A thing such that it is the unique property that is sometimes instantiated by human actions and does not obtain when the agent has no influence over the action, and ...”.) Ramsey sentences are particularly useful when a cluster of things are to be analysed at once—famously, an analysis of beliefs may need to mention other kinds of mental states (e.g. how they typically go together with desires to produce behaviour)—but we can make a statement of the roles of a range of mental states at once in non-mental vocabulary, if we replace the name of each mental property with a different variable. For example, if we were happy to take the notions of “action” and “content” for granted, then we might be happy with “There is one kind of thing, and another kind of thing, such that things of the first sort tend to cause rational behaviour together with things of the second sort, and both kinds of things can have propositional

---

the expression in the literature can be found in Lewis 2000.

<sup>2</sup> The expression is from Jackson 1998a. Jackson elsewhere refers to these as “placement problems” (Jackson 1994), and this terminology has been picked up by others e.g. Price 2004.

content, and...”, as a Ramsey sentence to simultaneously give us the roles of beliefs and desires.

The details of regimentation are not so important for the moment. (A canonical source for how to regiment is Lewis 1970.) Equipped with our specification of “theoretical roles” drawn from the platitudes, however we represent those roles, we are ready to take the second step in answering our demand for philosophical analysis. At the second stage, we look at our theory of the world to tell us what, if anything, plays the role so defined. Often the theory of the world we look to is formulated in a more restricted vocabulary than the theory with the vocabulary we initially used in saying what we wanted to analyse, though the theory of the world will hopefully have room for the vocabulary in our theoretical-role specification. We might use the deliverances of colour-science and psychology to see what deservers are available for our “colour” theoretical role, for example, or (one day!) use our theories of the working of the brain to look for deservers for the causal roles specified by our theoretical-role statement for beliefs, desires, and the rest.

Often the fit may not be exact: perhaps none of the structures in organic chemistry exactly play the role that “genes” are supposed to, or perhaps the brain does not quite have any kinds that behave exactly as we pre-theoretically suppose beliefs and desires do. Then the philosopher may have some more work to do in working out which of the available options is the “best deserver” to be the phenomenon under discussion—which property or thing satisfies the most of the role specified, or the most important aspects of the role satisfied. (The philosopher may also need to decide what to do if there are multiple equally good deservers, or when the best is still pretty terrible.) Some principles about “weighting” different platitudes will probably be employed, though this step of adjudication remains more of an art than a science. Once a best deserver has been identified, (or the appropriate excuses made) the philosopher’s job is complete: we will have an account of what colour, or causation, or free will, or genes *are*, and ideally an account in some more perspicuous terms than the ones in which we started. (Partial success is of course possible too: we may have a conclusion about what general sort of

thing our phenomenon is, or we may be able to rule out some alternative specific proposals even if a positive proposal for analysis remains elusive.)

Notice, however, that this “Canberra two-step” only helps us to explain how this conceptual analysis could be useful when we already have the answers required for the second step—when we already know what the world is like that we are going to place our phenomena in. So this will be of little help when we come to do fundamental metaphysics: when the Humean about causation and laws disagrees with a non-Humean about causation and laws, it is not a dispute about which of the commonly-agreed-upon entities best fills a pair of theoretical roles: someone e.g. who believes there is a primitive relation of singular causation believes in something in the world that the Humean does not. Likewise when we come to the debate between the dualist and the physicalist about the mind: it is not that the physicalist typically believes there are a complex of non-physical states with primitive intentionality associated with our brains, but is merely wondering whether our minds are those states or some physical states of our brain (perhaps identical to brain states, perhaps higher-order functionally defined states). No—the dualist believes in a range of states that the physicalist does not, and it is not very clear how arguing over the concept of “belief” is going to settle this issue.

So there are some metaphysical questions that will not be susceptible to “location” solutions, without some further work to sort out what there is in the world for the phenomenon in question to be identified with. There will be many more questions where it will be to some extent controversial whether they are mere “location” questions. Many people treat debates about personal-identity over time as location debates—there are sequences of minds (or stages of minds) that are psychologically continuous, sequences of bodies that are physically continuous, various patterns of informational and causal dependence, attitudes and behaviour of groups of bodies and psychologies... Which, if any, of these sequences defined in these terms are sequences of the same person? But others think we need some additional metaphysical resources if we are to identify the best candidates to be the relation of identity over time in this world. Perhaps we need haecceities or essences, connected in some way to the distribution of qualities at times so

as to yield distinguished sequences. Perhaps we need a new relation of non-reductive constitution that people can stand in to bodies and psychologies. Maybe we need changeless souls to preserve us in a Heraclitean world of change and decay, maybe we need irreducible substance-sortals... or whatever. Can we derail the location program whenever we have a particular metaphysical problem to solve just by making some sort of *prima facie* case that we need a new fundamental metaphysical commitment? If so, then conceptual analysis will be of little use in most of metaphysics, except perhaps to resolve debates between people who are sufficiently committed to a particular ontology that they will not revise it because of the metaphysical issue in question.

This is a particular worry if solving location problems (and doing the philosophy of language necessary to get the correct machinery in place) is all that philosophy can do, or should do. For if the second part of the two-step is something that we should leave to scientists, for example, then it does not look like there is much left for philosophers to do in metaphysics. This would be particularly bad news, given how bad many scientists are at metaphysics—they often don't know how to do it, and don't care. Which is sometimes fair enough—not everyone can do everything—but it does suggest there's a range of important issues that won't get properly addressed by anybody. Presumably what we should think is that philosophers should not limit themselves to just being in the business of solving location problems—this is a conception of philosophy that could eventually be as stifling as the Wittgensteinian conception of philosophy as only diagnosing and dissolving confusions resulting from misunderstanding language, confusions which according to Wittgenstein and his followers were often due to philosophers in the first place.

Fortunately for the Canberra plan, and arguably fortunately for metaphysics, the project of assembling "platitudes", and seeing what they jointly say about the subject in question, can be defended as valuable even when there is an issue of the second kind to be resolved. This paper explores three ways this method can assist in the second project. I do not know whether all of these usages should count as "conceptual analysis"—perhaps we will need to solve the "location problem" for the concept of conceptual analysis itself

to tell! What I do hope to show, though, is that much of the method is of a kind with the method that Jackson, Lewis etc. take to be what we are doing when we are doing conceptual analysis proper. In doing so, I will have pointed out more options for the method, and helped to solve the puzzle I began with: that is, solving the problem of what philosophers could justifiably be doing when they employ armchair methods to solve metaphysical problems. The significance of this goes beyond metaphysics, in fact—there are many areas of theoretical inquiry where something like the methods of conceptual analysis seem to be at play, for example where people engage in thought experiments. Often what people are aiming for is not just information about our descriptive capacities or our concepts, but rather information about the world. (Thought experiments in physics, for example, are employing in an inquiry about what the physical world is like, not an inquiry primarily about how we represent it.) A more satisfying account of how conceptual analysis could help with such questions would be welcome.

Before I outline the three ways I think a platitudes analysis can be helpful, I think it is worth discussing a few issues about what the platitudes analysis itself involves—or at least what it involves for my purposes. Hopefully this will not only serve to clarify what I have in mind, but to sidestep, or address, common objections raised to the Canberra plan of conceptual analysis.

### *The Platitudes Analysis: Filling in the Details*

"Platitudes" is an expression that is to some extent a placeholder—it is a name for those claims or principles we use as the raw material for the process I outlined. But before I go on, I should say something more about what we might think these platitudes are. This is not so much because I want to defend a particular answer—indeed, there may be no unique answer, especially if the platitudes analysis might be employed for different purposes on different occasions—but for two other reasons. One is to give a sense of the range of different options there seem to be here, since I have come across different conceptions being employed by different people, and I think there may be some risk of people talking past each other on occasion. The other is somewhat related. Sometimes

objections are made to the method of platitudes analysis on the basis of certain specific conceptions of what the platitudes are and how we are to tell what they are: and while those objections may be good as far as they go, if it is true that the structure of the account can be preserved with any of several conceptions of the platitudes plugged into it, then we should be careful not to reject the framework altogether simply because of concerns about one way of fleshing it out.

So, what are the platitudes? Or rather, a better question might be what the platitudes are for a certain term or concept: e.g., when we are analysing “belief” (either the word or a mental analogue), what sentences or other representations about belief ought to go in our list of platitudes? (Or perhaps when we are analysing beliefs what we should ask for are all the platitudes concerning mental terms—not just beliefs, but desires, intentions, actions, perceptions etc.) After all, what we need for any given occasion we want to find out the “platitudes about *X*” (or the platitudes concerning *X*, or whatever)—we rarely, if ever, need to know what the platitudes are *simpliciter*. One version of this question is the question of what kind of thing a platitude is: is it a belief, or a sentence, or a proposition, or something else? If we are analysing a word or expression, such as “belief” or “free will” or “law of nature”, then I suppose it is natural to think that platitudes concerning these will be some of the sentences involving these expressions (or perhaps something expressed by these sentences—if we think the platitudes about laws of nature should be the same in English and French, for example). If we are analysing concepts, like the concept of belief, or the concept of a law of nature, or whatever, then the platitudes might be beliefs. Or in either case we might say they are propositions—but if they are propositions, they will presumably be structured propositions, given how they are manipulated (see below). If we are talking about analysing some phenomenon, on the other hand: colour or causation as opposed to “colour” or “causation”, then presumably we might have in mind any of the three tasks best served by looking at interpreted sentences, beliefs, or structured propositions respectively. Rather than being careful about this, let me talk about platitudes as if they are sentences, and when I talk about something figuring in the platitudes, let me be taken to be talking about words or expressions. (So were I to say, e.g. that assertion appears in some of the platitudes about

truth, the initial gloss should be that “assertion” appears in some of the sentences including “truth” that count as platitudes for truth.)

One story, perhaps one of the earliest ones, says that the platitudes about some *X* are all the sentences talking about *X* that we take to be true. (*X* may be a single thing, or it may be a family of things—the platitudes about the mental, for example, presumably include platitudes that include words like “belief”, “desire”, “thinks”, “is conscious of” etc.) Something like this seems to be what is going on in Lewis’s seminal “How to Define Theoretical Terms” (Lewis 1970), where the Ramsey sentence involved in defining a set of terms we want to analyse (the T-terms) is a “T-postulate” that contains all of what our theory tells us using the T-terms. A related story, which is closely aligned with the Lewis 1970 one, is that we identify some theory containing the *X* vocabulary, and it is only the sentences that make up that theory which are the platitudes for the *X*s. Which theory? Here we have a range of options. For technical terms with a limited community of users, we may be able to point to some defining theory. For other terms, we may be able to locate a “folk” theory—as some people hope to find a folk psychology, or a folk physics, or a folk biological taxonomy. (Lewis 1972 p 257 says that mental terms are defined by “common-sense psychology”.) The theory we use to fix the platitudes may or may not be one we currently accept—for instance we could take “tree” to be implicitly defined by a folk biology which we take to be mistaken in places, or if we want to know what the platitudes are about caloric fluid we may take some theory or theories of caloric fluid to provide the platitudes, even though we now reject such theories.<sup>3</sup>

At the other extreme, we might be very restrictive about which of the sentences we accept should count as platitudes. For example, we might take a platitude about *X*s as something someone believes on pain of being linguistically incompetent with the application of the *X* terms. (Or we may require that someone believes the relevant platitudes as a necessary conditions for possessing the *X*-concepts, to formulate the equivalent constraint on the mentalistic conception of platitudes). Formulated this way, platitudes are treated rather

---

<sup>3</sup> There are other variants along these lines—instead of listing all of the relevant sentences we take to be true, or all the relevant sentences in a theory, we may instead wish to list a set of sentences which are some privileged organization of the relevant sentences (e.g. a set that jointly entails all of what we believe).

like traditional analytic/conceptual truths were supposed to be: someone's competence with "bachelor" could be called into question if they didn't believe that all bachelors were unmarried.

Formulated this way, traditional suspicion of analytic/conceptual truths translates into suspicion that there are any such platitudes. We might doubt that there is *any* error that *must* be attributable to a fault in linguistic competence rather than a mistake of fact, or we might think that very little is required in the way of belief for someone to be linguistically competent with an expression. Or we may think that some expressions have some such constraints, but many do not. (What do we have to believe in order to competently classify something as falling in the extension of "cat"?)

Many people are not explicit about what they count as a platitude. But I have a sense that there is clustering at two places on the spectrum between counting every sentence containing the vocabulary that we accept, at the one end, and accepting only the non-negotiable, at the other. One place is close to the non-negotiable analytic truths end. Platitudes are truly central to what it is to be an *X*. (Call this the "Strict conception").<sup>4</sup> There are two, or possibly three, respects in which people may vary from a fairly traditional insistence on analytic truths. The first is that platitudes are not quite non-negotiable: there may be conflicts between platitudes, and in this case (and maybe only in this case), we need not insist that a phenomenon satisfy all the *X*-platitudes in order to count as an *X*. One example of this might be the naive theory of truth—an unrestricted T-scheme together with a straightforward naive theory of quotation enables us to prove a Liar sentence is both true and not-true: most people think this shows one of these platitudes about truth and quotation should be not accepted as-is. The second is that we may wish to treat some sentences as *X*-platitudes, not because competent users accept them, but because competent users accept that they would be true *were there Xs* (or *X* was instantiated if it is a property, or obtained if it is a state, or whatever). It may be a

---

<sup>4</sup> I take an example of this to be the use of "platitudes" in Smith 1994: see pp 29-31. A related usage is that of Crispin Wright in, e.g. Wright 1998: Wright is not a Canberra Planner, but his usage is obviously related (and he tells me the relation is ontogenetic as well as phylogenetic). Examples could be easily multiplied.

platitude about caloric fluid that it is transferred if and only if there is heat transfer; or a platitude about God that s/he created the universe. But we may insist, not that competent users accept these claims (assuming they are genuine platitudes), but only that they recognize that they would be true *if there was such a thing as caloric, or if God existed*. A third variation can be in the kinds of claims one takes to be platitudes. Influenced by Kripke, Putnam and Burge (Kripke 1980, Putnam 1975, Burge 1979), platitudes for Xs of the form “being a causal origin of our use of the expression ‘X’”, or “being what the experts intend to refer to with their use of ‘X’”, or “is a natural kind that *actually* thus-and-sos” are the sorts of platitudes that we may find on our list for some terms.<sup>5</sup> Whether this variation is a departure from the traditional conception of analytic truths, or merely a variation within the bounds of that tradition, is not clear.

On this strict conception, platitudes define our subject matter in a straightforward way—to disagree with a platitude is to make a conceptual error or alternatively to change the subject. This gives us relatively little flexibility, and the platitudes, if they are to be this central, may not be informative enough to give us a unique fix on our subject matter. Of course, some of those who take platitudes to be this unyielding will be happy to indulge in changing the subject, providing the subject is not changed too far: a surprising example is Frank Jackson (1998a p 45), who claims that our concept of identity of persons over time requires an irreducible further fact, beyond physical or psychological continuity or similar, and in fact requires something “never instantiated”. Jackson is committed to the view that there are no people who exist at more than one time, given the meaning “people who exist at more than one time” has. (Since I doubt physical systems like us could be instantaneous people, I presume the thing to say on Jackson’s view is that there are no people. I wonder whether there is mentality and agency or any of those things either, given what Jackson requires for there to be people?) Unsurprisingly, Jackson advocates that we use a replacement concept—a “different and ‘nearby’ conception”, to use his words. In the language of platitudes, Jackson believes there are platitudes about personal identity (or about personhood and time, perhaps), that a relation

---

<sup>5</sup> This would not be the influence intended by these authors, if we took platitudes to capture the meaning of theoretical terms, of course.

of identity over time for persons must satisfy – and these platitudes are sufficiently non-negotiable that a relation that does not satisfy them is not identity through time for persons. Occasionally I have heard talk of “changing” or “adjusting” or “revising” the platitudes associated with a given term—I am not sure whether this is talk about choosing to use a different expression with a different sense (but which is spelled and pronounced the same), or whether it presupposes that we can change the meaning of an expression in our language just by making a decision about it (perhaps this is true for some technical terms, but I doubt philosophers usually have sufficient linguistic authority to change our language or our conceptual scheme in this way), or perhaps this is talk of something else.<sup>6</sup> No doubt the speakers are not always clear what they have in mind either.

There is an alternative conception—I don’t have a group to pin it on yet (maybe it is idiosyncratic). This alternative conception falls closer to the “kitchen sink” end. Not everything we believe (or that is claimed by a defining theory) need go into the mix, but quite a bit does. Going along with this, platitudes are conceived of as being more fallible, and individually the X-platitudes are less like necessary conditions to be an X. The place where this kind of conception is most in evidence is in analytic functionalist treatments of folk psychology. Many of the psychological generalisations we are inclined to believe will, no doubt, be shown to have counterexamples, or not to apply straightforwardly to unusual cases.

The progress of cognitive psychology should show us that there are things still to be learned about the behaviour of beliefs, desires, intentions, emotions, and the rest. Not every generalisation we believe about ordinary psychology will be vindicated, and it may be fruitless to try to identify any particular generalisation that we expect will be unrevised. But according to analytic functionalism it is still the case that whatever beliefs, desires etc. are, they satisfy many or most of the generalisations about their

---

<sup>6</sup> One thing it might be talk of is only changing what we *take* to be the platitudes associated with a term. While strictly speaking this would not be changing the platitudes but only changing our beliefs about them, English does permit this way of speaking (perhaps it is metaphorical). It would be similar to the way of speaking in which we can appropriately say that nineteenth century physicists changed the electron several times, or a modern archeological discovery changes the ancient Egyptians from an insular culture to one embedded in very long range trading networks.

behaviour provided by folk psychology. Nevertheless, if the only “mental” states mature psychology were to deliver us satisfied completely different generalisations and claims about their interrelations than the folk psychological theory about mental states, then that would show that there were no such states as the familiar beliefs, desires, hopes, fears, pains, emotions, and all the rest.

Once a much more generous conception of platitudes is embraced, then we need to say what the limits on this generosity are. If we allow every claim we believe about a given topic to be a platitude, then we are back with something like the Lewis 1970 view. (Lewis was talking about theories rather than sets of beliefs, but we can treat a person’s beliefs as a special case of a theory, albeit an often inarticulate and messy one.) Even at this extreme, there are still different ways to go. When I am considering what the platitudes are about personal identity or free will, I might include all of what *I* believe about the topic (or at least as much of that as I have straightforward access to.) Or maybe I should impose a requirement of common belief—perhaps, even at its most generous, the platitudes are only the things that are believed in common in my community. Or perhaps we should be even more restricted, and include only the things that are not only believed throughout our community, but about which it is common knowledge that they are so believed.

If we are to be less individualistic, and decide on what are platitudes on the basis of shared beliefs or shared beliefs that it is common knowledge are shared, then the question also arises at this point about which community is the one that we should be concerned with. All of humanity, or all of contemporary humanity? People with a similar culture, or similar patterns of engagement with the world? People who think as I do (which may mean, in effect, that what we believe is just what I believe)? I expect our verdicts here may vary from case to case. In the case of platitudes employing technical terms, or terms which began as technical terms, such as “electron” or “supervenience”, or “gene” or “entropy”, we may be primarily interested in the beliefs of experts, somehow identified. In the case of “folk” notions, like “thinking” or “heat” or “artefact”, we may think the community needs to be selected more inclusively.

One limitation on the size of the relevant community may be linguistic. I have been talking loosely of platitudes being given by what is believed, but strictly speaking platitudes are going to be sentences (interpreted sentences). Given this, we may wish to restrict our communities to those sharing a common language. Those who speak no English are unlikely to accept the sentence “Some beliefs are stronger than others”—or any other English sentence, for that matter. Of course there is linguistic variation among English speakers as well—we will get a very strange picture of what it is to “luck out”, for example, if we do not restrict our platitude gathering to one or other of the patterns of use of that expression. (North Americans use “luck out” to describe situations in which one is lucky, British and Antipodean speakers use it to describe situations where one is out of luck.)

Going to what is in common in a linguistic community is one way to narrow down the platitudes to something less generous than just including any old thing I happen to believe on the topic. Another way is to include among the “platitudes” only the things that I take to be clearly true, or obvious, or certain about the relevant topic. When considering the platitudes about possible worlds, for example, the claim that a world has some things true according to it is plausibly platitudinous: that there is a world that contains more individuals than sets is not platitudinous, even though it is something I am tempted to believe. We could explain this by saying that the second belief is idiosyncratic (which would invoke some standard of commonly shared belief), or we could explain this by saying that the second belief is tentative, not something I take to be obvious, and something that I find myself inclined to think requires justification in terms of more evident claims about modality, possible worlds, and so on.

If we raise the standard for something to count as a platitude in this way, then “platitudes” become more like traditional philosophers’ “intuitions”: though notice that clearly true or obvious claims need not be particularly a priori or philosophical: it is a platitude of human physiology that one’s nose is typically smaller than one’s foot (albeit that it is platitudinous in the ordinary as well as the technical sense): but we need not

suppose that this claim about noses and feet is distinctively philosophical, nor discoverable through some privileged faculty, nor a conceptual truth, nor any of the other things that some have been tempted to think about typical philosophical “intuitions”.

Another constraint we might try to put on platitudes involves their origin. There are some things that one has a tendency to believe for specific reasons, and other things that one is inclined to take for granted. This can vary from context to context, of course, and it is also true that what is taken for granted by somebody may be quite specialised, technical, and/or esoteric information: what a particle physicist takes for granted about quarks will not be much like an unquestioned dogma of everyday life. Perhaps when people are assembling platitudes about the mind, or causation, or moral rightness, or whatever else is the target of their philosophical investigation, some sensitivity to whether a claim is taken for granted or seems eminently contestable may play a role in the sorting procedure.

Once we have criteria that come in degrees, such as degree of common belief or common knowledge, or degree of certainty, or degree to which something is taken for granted, we are provided with a way of assigning “weightings” to different platitudes, in proportion to the relevant quantity. Exactly how to weight platitudes is a vexed question, especially if more than one factor that comes in degrees makes for platitude status. Questions about how platitudes are weighted are of particular importance when platitudes conflict, since people often think that one platitude can trump another (or one group of platitudes can trump another group) if one is much more important or weighted much more highly than the other. I have the impression that those engaged in the Canberra plan do make judgments about relative weightings of platitudes, and these make a difference in decisions about what to reject when there is a conflict (or when nothing in the world precisely satisfies the original list of platitudes), but I am less confident about what explains these judgements.

In making the above remarks, in part I am trying to characterise what is going on when those who are generous about “platitudes” sort platitudes from the totality of what they

believe about a topic, or the totality of what a theory they accept says about a topic. The characterisation has not been altogether precise, in part to give a feel for a range of variation, and in part because I am not entirely clear myself what unarticulated principles are being deployed. But one of the things I mean to be doing in providing the above sketch is just to indicate an end-point on a continuum of positions. At one end is the strict conception of platitudes as more-or-less traditional analytic truths, and at the other end are those who include most of what they take to be true, or take to be common belief, or that is pretty clear or certain or obvious about a given topic. Individual people who find it useful to assemble platitudes as part of philosophical investigation may fall closer to one end or the other of this spectrum (or may even fall in different places on it on different occasions, or when dealing with different subjects). Or I suppose for many people who engage in this project it might be difficult to tell where they fall on this spectrum, or there may be no very determinate fact of the matter.

### *What to Do with the Platitudes*

Whatever the account of what platitudes are and how we collect them, eventually we come to have a collection of sentences which we take to express platitudes about a particular subject matter (let us assume for convenience these are all in a single language). We can then use this collection to discover things that might not have been obvious from our implicit grasp of the platitudes taken in isolation. For a start, it may only be after this process that we come to see things that require entertaining several platitudes at once. One example is illustrated by an argument for compatibilism in the free will debate. Initially, we may have the intuition that our actions being determined by past events beyond our control is inimical to free will. We may also have the intuition, though, that if our actions were random that would also not be a case of free will. When we put those together with the suggestion that there isn't much more (or maybe anything more) to something's being random than its not being determined by any previous state of the world, we might be inclined to look more favourably on the idea that some sorts of determination might be compatible with freely willed action. Each principle (the one

about determinism precluding free will, the one about randomness precluding free will, and the one about randomness following from a lack of determination) has some initial plausibility—each has some call to be called a platitude, at least on suitably generous conceptions of platitudes. But it is only when we put all three together, or put all three together with the “platitude” that some of our actions are freely willed, that we see that there might be good reason to revise our principle about the incompatibility of free will and determinism. Of course, whether the adoption of compatibilism is the *right* response to the discovery of this tension is controversial. But I am merely using this as an example of how one might think some important philosophical insight may come from considering together thoughts that are individually platitudes, or at least seem to have some claim to that status (especially if we have a relatively generous conception of platitudes). Note that this example is a case where the compatibilist at least can be seen as arguing that one of the platitudes is, in the end, incorrect—so it sits less comfortably with the picture of platitudes as core conceptual truths about a subject matter.

In the account of solving “location problems” I gave in the opening section, I more or less took it as read that we would have a theory available to tell us what the deservers for the platitudes are. But there should be more to say about how this part of the task is carried out. Sometimes it will be controversial what the world is like. Sometimes the only theories of the relevant bit of the world we will have available will employ the vocabulary or concepts under investigation. And sometimes, perhaps often, our theory of how the world is will not deliver up any things that precisely play the family of roles given to us by a group of platitudes. I expect people employing platitudes analyses can fruitfully disagree about what to do in these situations.

There is one final dimension of possible difference in approaches to platitudes analyses that I want to mention. This range of differences centres on our justification for thinking that the “deservers” identified by this method really are the referents of the terms or concepts that appear in our network of platitudes. A view traditionally associated with the platitudes analysis is that the platitudes associated with a group of expressions together give us the *meaning* of those expressions, and so it is *analytic*

that any group of things that jointly satisfy (or best satisfy) those platitudes are the referents of the given expressions. Many of the proponents of a platitude approach are descriptivists about the meanings of theoretical expressions, at least: Carnap developed the method of employing Carnap sentences to capture the meanings of theoretical terms in part as a response to Quine's attack on analyticity (Carnap 1963 958-966), so that the platitudes analysis was capturing the meaning was important to him. Lewis suggests that this method gives us *analyses* of the relevant expressions, where this seems to be a claim that it serves up analytic truths. (see e.g. Lewis 1989, and see Nolan 2005 213-227). And Jackson is a well-known defender of descriptivism about the meanings of a range of expressions (see e.g. Jackson 1998b) and his defence of analytic functionalism about the mind makes clear that the folk-psychological platitudes are supposed to deliver the meanings of mental vocabulary, as well as enabling us to determine the extensions of the relevant predicates and referents of the relevant property names. (Braddon-Mitchell and Jackson 1996 ix, 46).

Descriptivism about theoretical terms may have much to recommend it. But note that for several of the core purposes that these sorts of analyses might be used for, it is not required. Suppose you thought that names of properties associated with e.g. natural kind terms were directly referential. For instance, let us suppose the expression "Being a Whale" took as its semantic value simply the property of being a whale. You assemble the "platitudes" about whales: these might be the platitudes among experts, if you are particularly interested in the expert usage, or they might be platitudes in the community at large. Suppose, in any case, the platitudes you assemble make claims that you at least believe are true, or true for the most part. (Perhaps you defer to whale experts and believe what they tell you about whales.) Suppose, when you put the platitudes together in the appropriate way, they tell you that Being a Whale is the unique property such that it PHIs, and its instances behave *thus-and-so*. (Swim in the ocean, are such that some of the instances eat plankton, are all such that they give birth to their young live, etc.) You might then use this information to go looking for the property such that it PHIs and its instances behave *thus-and-so*. Suppose you find one, describable in alternative vocabulary. (Let us suppose that the property of being an animal with DNA of such-and-

such composition is, indeed, the unique property such that it PHIs and its instances are thus-and-so.) It seems that you are in a position, given what you know, to infer that Being a Whale is identical to the property of Being an Animal With DNA of Such-and-Such Arrangement.<sup>7</sup> And it seems that you are able to do this without making any assumptions about descriptivism vs direct reference for the expression “Being a Whale”. If you believe the platitudes; know that they imply that Being a Whale is the unique property such that it PHIs and its instances behave *thus-and-so*; and know that being an animal with certain DNA is a unique property such that... , then that is all you need for the inference.

At least when the platitudes employed are ones you believe, or when you believe that they are largely correct, you will have reason to think that if they have a good enough group of things satisfying them, that those things are the things referred to. This is so even if you have no idea how the expressions get their referents. You might be torn between descriptivism and some sort of direct reference (or some third option) for mental vocabulary, perhaps even because you are not sure whether the expressions behave more like natural kind terms or like functional vocabulary, because you are not sure whether e.g. beliefs form a natural kind or only share a functional specification. You can use the platitudes about mental states to as a guide to what the property of being a belief amounts to, and then see what sort of property, if any, (functional, natural kind, other) does well enough satisfying that specification, and *then* hypothesise about how that property is associated with the word. As Carnap pointed out, it is your theory that commits you to the corresponding Ramsey and Carnap sentences, not anything else. The further claim that the associated Ramsey or Carnap sentences have much to do with the meaning of the terms in your theory is an extra step, of course, but producing a specification of roles and hunting for deservers does not need this step, if all one is interested in is what in fact the terms refer to, if anything.

So taking the network of platitudes to capture some or all of the meanings of the relevant

---

<sup>7</sup> This example is merely illustrative—some might think that it will turn out to be the property of having a certain ancestor, or will involve the possibility of interbreeding, or something else. This is not intended as a serious conjecture in the philosophy of biology.

expressions is an optional add-on, it seems to me. It does bring some benefits, of course. One advantage is that it would enable us to return verdicts about merely possible cases as well as actual cases. If some hypothesis is entailed by ordinary analytic truths, it will be necessary. So if we are interested, for example, not just in the question of what beliefs and desires happen to be, but also what the limits are on what they could possibly be, then if we can take a group of insights into beliefs and desires to be analytic then we will have much more guidance on the broader question. A related advantage is that if the conclusions one reaches are supposed to be analytic, then they can be tested more easily by considering merely possible cases. If the claim that people often desire to do what they believe is right is merely a commonplace, then it may tell us very little about whether there are possible creatures that do not have any desire to do what they believe is right. But if it is a putative analytic truth about “right” (or the combination of “right”, “belief”, “desire” etc.), then we can test to see whether it has that status by considering the intelligibility of a thought experiment where we describe such creatures. Finally, if we both thought that platitudes were analytic, *and* that analytic truths were implicitly grasped by those competent with the relevant concepts, then we might be more confident that purely armchair methods will enable us to discover them. If I am confident that I competently deploy the concept of causation (and competently use the English word “cause”), I can be confident that I already know the platitudes about causation, at least implicitly, and all I need to do is whatever self-examination is required to make them explicit. If they are not analytic, I need some other reason to be confident that I can come up with them merely through reflection. Likewise for platitudes about truth, moral rightness, colour, personal identity, etc. If I had good reason to believe both that the platitudes were analytic and that the analyticities were all things I implicitly knew, I could have methodological confidence in armchair philosophy that might otherwise be harder to come across.

Despite this appeal, some will not want to take the extra step of identifying the meaning this way. They might, for example, be convinced that this sort of descriptivism about theoretical terms is ultimately mistaken. Perhaps they are convinced that the world can be sufficiently surprising so that competent language users cannot tell in advance that

many of the platitudes are correct (maybe science could have determined that electrons are a kind of German sausage after all, if the evidence had been sufficiently bizarre?), or perhaps they subscribe to a philosophy of language according to which meanings are determined by word-world connections that provide for very little analytic truth. Or they may even be undecided, or wish to stay neutral—they may have enough challenges of their own to worry about, without taking on a controversial commitment to descriptivism. The benefits for metaphysical inquiry I will be suggesting should appeal even to those who reject, or are suspicious of, any straightforward connection between platitudes and meaning.

### *How the Platitudes Can Help: The First Way*

The difficulty, it seems, is that when we are doing fundamental metaphysics we are trying to work out what the world is really like, but all that the method of putting together the platitudes gives us, according to Jackson (& co) is something else—it gives us information about a term or concept that, *when put together with a story about how the world really is*, tells us what that term or concept applies to. The first way I want to mention in which putting the platitudes together can help provide answers to metaphysical questions is a relatively familiar sort. Getting clear on our concepts and meanings can help us disentangle different options, it can give us information about what is conceptually possible, and it can help sharpen the questions we then attempt to answer.

These are traditional tasks for conceptual analysis, and it may explain some of the armchair use of platitudes in metaphysics—especially the use of deductive reasoning to show things that are not obvious to competent users of the concepts. *If* it could be shown that omnipotence and omnibenevolence were mutually inconsistent, we would not need to investigate the world to see if there was an "omni-god" as conceived by some strands of traditional theism. *If* the occurrence of change was inconsistent with there being a determinate future, that might drastically constrain our options for a metaphysics of time. *If* we have to suppose that mental states sometimes are caused by perception and cause behavior, we can rule out epiphenomenalism about mental states (at least across-the-

board epiphenomenalism), without needing to investigate whether there are any non-physical epiphenomenal states.

In a similar vein, becoming aware of slight variations in meaning and views can be valuable preparation for doing good work, in metaphysics and elsewhere, and it is something that philosophers in the analytic tradition typically pride themselves on. Whether it is pointing out that the questions of what a cause is and what is relevant to a causal explanation are different questions; or that the hypothesis that "everyone always acts selfishly" has a range of different meanings ranging from the implausible to the possibly quotidian; or that "truth" for sentences and "truth" for propositions would be quite different, though presumably connected; distinguishing options is something that can to some extent be done usefully in the armchair, on the basis of little more than a grasp of the concepts or the meanings of the relevant terms.

Another thing that collecting the platitudes can do is help define our subject matter (see Jackson 1998a p 30-31). If you ask someone off the street whether there are any properties, and if so what are they, my guess is that person will have little idea what you are asking (and this is true even after we clear things up a bit—by making it clear we are not talking about real estate, we are not talking about whether people sometimes use abstract nouns, and maybe even after we even give some examples of putative properties like colours and shapes, etc.) One thing that collecting platitudes can help us do is to get a clearer handle on what sort of question we are asking. Would there have to be things that were multiply located in order for there to be something that counted as a property? Or is that just an optional add-on (which e.g. we may want to include to help answer the question of what it is for a thing to instantiate a property, or we may want to reject when the time comes). Do properties have to be causes, or effects? Do the properties had by distinct particulars have to be numerically identical, or not? Some of these matters may be able to be settled by old-fashioned conceptual analysis—some claims about properties may be so central that to deny them is tantamount to either contradicting yourself or changing the subject.

Removing conceptual confusions about possible answers, distinguishing possible answers that may otherwise be run together, and getting clear on what the metaphysical questions are, are all fruits offered by traditional conceptual analysis. They are the sorts of things that should be valuable across the board in almost any inquiry about the world, and not just metaphysical inquiry. They may be particularly useful for the kinds of questions that philosophers address, of course, because those questions can be very general, and lack a generally agreed methodology to answer. (And other disciplines often have more clearly defined questions to address, or so it seems to me—though I suspect coming up with the right question to ask is important in virtually every field, even fields with better defined methodologies for answering them.)

I suspect the platitudes approach will deliver less along these lines than traditional conceptual analysis was supposed to. This is because the platitudes approach is likely to yield a cluster of descriptions, any single one of which is likely to be negotiable if a candidate does well enough in satisfying the others. (So we may need to revise our off-hand opinion about how many objects there have been in the room in the last week, if we accept the existence of temporal parts or sub-atomic objects, and we might need to accept that things come to be at different temperatures when the same amount of heat is applied to them, or to give up other common-sense views.) We might have to live with a few thought experiments delivering counter-intuitive results, or a few commonplaces being overturned. This can be seen either as bullet-biting, or as making exciting metaphysical discoveries, according to taste. This makes things more messy—bullet biting can seem ad hoc, and questions like the relative importance of different platitudes are difficult to settle and potentially endlessly controversial—especially given the human tendency for proponents of a theory to weight highly the respects in which a theory does well, and minimise or disregard the respects in which a theory does poorly.

On the other hand, we do not get to choose whether or not our methods are going to need to be messy. I would like a sharp, clean method in conceptual analysis as much as the next person—a way of showing that a counter-example is indeed fatal to a view, or can in good conscience be resisted provided some definite set of criteria are met. Perhaps one

day things will be like that—Leibniz's dream that philosophical disputes could be settled by mutually agreed upon precise algorithms might be down the track somewhere—but I'm not holding my breath, and I don't recommend anyone else does either. I am convinced by some of the fairly standard examples that definitions can be widely accepted and core platitudes can be taken to be extremely important, and yet turn out to be mistaken. To take an example from the history of physics, everyone accepted that the definition of momentum was that it was a quantity equivalent to mass times velocity. This wasn't even taken to just be a well-confirmed fact—it was taken to be a *definition*. One might argue whether momentum was constant in all interactions, but not, one would have thought, whether it was directly proportional to mass times velocity. Relativistic physics convinced us that momentum was not quite proportional to mass times velocity, however. In fact, the quantity of momentum as relativity conceives of it can be indefinitely different to the product of an object's mass and velocity (at least its rest mass and velocity). Instead, we think that there was enough about the Newtonians' use of the word "momentum" that this quantity is what they were talking about—it is conserved in reactions, is roughly proportional to mass times velocity at low speeds, it explains the observations that led us to use the expression "momentum" in physics in a fairly straightforward way, and so on. The lesson I want to draw from this case here is that we are likely to overestimate the importance of any single platitude (such as the official "definition" of momentum, which was a core platitude). And we should keep in mind that something may turn out to be close enough to the set of constraints that are in place in unexpected ways. On the other hand, while some may use the momentum case to argue that no combination of platitudes constrains what "momentum" may apply to, I do not think it shows this—momentum couldn't have turned out to be a kind of cat, or be identical to the temperature 24C. (To be old-fashioned, I'm even prepared to suggest that these claims are analytic—you can be in a position to know that momentum isn't a kind of cat simply in virtue of understanding expressions like "momentum" and "cat".)

So I am suspicious when someone claims to solve a traditional philosophical problem—the nature of free will, how objects continue in existence through time, the nature of causation—through the platitudes analysis (or other sorts of conceptual analysis) alone.

Our terms and concepts probably have too much slack to settle these questions. But this is not to say a platitudes analysis is not helpful in these ways—I think there are useful contributions that conceptual analysis in general, and the platitudes analysis in particular, can make to investigations of the world—and crucially, these are contributions that can be recognized independent of knowing the details of what is in fact the case.

Recognizing analytic constraints on what descriptions of the world are coherent is a traditional conception of philosophy, or at least part of it. That conceptual analysis can provide useful distinctions and help us to get clear on our subject matter are not exactly new ideas, either. Together these enable the platitudes account to help play the "under-labourer" role in investigations of the world, something like the role that Locke suggests for his philosophy. This applies to metaphysics as much as to other inquiries about the world—the platitudes analysis can stand to metaphysics as it stands to other inquiries, including many outside the current disciplinary bounds of philosophy.

### ***The Second Way***

The above virtue of the platitudes analysis is one that we should expect defences of conceptual analysis more generally to claim for their enterprises. The other two benefits from a platitudes analysis are less like the sorts of advantages that might be expected from traditional conceptual analysis. This might be an advantage of the platitudes account over some traditional forms of conceptual analysis: it may also be a reason to think of the platitudes analysis as not being a kind of conceptual analysis at all. One feature both of these benefits have is that the support they provide for a view is typically *a posteriori*—which is not part of the original picture of conceptual analysis!

Assembling platitudes about a subject matter (whether causation, or free will, or genes or acids or whatever) often results in providing a statement of folk theory—or at least the less speculative parts of folk theory. For example, if we were trying to assemble the platitudes for "tree" (in the sense of a kind of vegetation), we would mention trees being plants, trees usually having leaves, trees being largely made of wood, etc. etc. Maybe some part of this story would be analytic, or definitional, but much of it would go beyond

that. Articulating widely shared implicit beliefs employing a concept (or stated using a term), together with drawing consequences from the body as a whole, will probably reveal a theory that will be incomplete in all sorts of ways but will likely articulate a set of commitments we may in fact already have, or at least that we are inclined to have. Or when it is the platitudes of some group of specialists we are investigating, we may get a set of commitments that our experts share and we are inclined to defer to. (Or on occasion someone may be both a conceptual investigator and a "specialist" in the relevant sense, as when a theoretical biologist tries to articulate the concept of a "gene" in population biology, or a philosopher tries to do a conceptual analysis of what it is for something to be a proposition or a possible world.)

Assembling platitudes and seeing how they fit together is obviously not the only way to discover what we believe. Often we can tell a lot about what our views are just by introspection, and there are other ways which also require little more than an armchair—I often discover things about what I believe in conversation, or by reading or listening to someone else articulating what they believe, and trying to work out whether I agree. Coming up with my "pre-theoretic" opinions about a topic may be done in a similar way, though there is always the problem of contaminating one's pre-theoretic views by paying them too much attention—it is hard to avoid slipping into theorising, and hard to avoid attributing to one's past self the beliefs one thinks one *ought* to have had. If my only task was to come up with what I believe, coming up with platitudes might seem a fairly inefficient way of doing it. But there are several things platitudes can be relevant for in this area, some of which are harder to do by mere introspection.

In metaphysics as elsewhere, we form a community of inquiry. (Less so than in many areas of research, I'll grant.) As well as working out what I believe, I often want to know what *we* believe: *we* might be the philosophical community at large (or some subsection, as when I try to work out what I think the orthodoxy is on a given issue), or it might be some sub-group who is sympathetic with my views about some issue (what do the realists believe, or what do the objectivists in ethics believe?), or it might be narrower again (what consensus is there in my current conversation, or among the people I talk

with about a certain question?).<sup>8</sup> Again, asking people or reading what they have written on a topic can often serve very well for answering these questions: but the kind of investigation to find out what there is agreement about will be much the same as investigation into what the platitudes are for the relevant group, especially if we ask about relative certainty and relative unrevisability.

Another thing we might be interested in often goes under the name of "common knowledge": what are the things that are not only believed by a group, but that the group believes each other to believe, and believes each other to be aware of this in turn, and whatever else distinguishes common knowledge from knowledge that happens to be had by each of the members of a group (or most of the members of the group). The only problem with the label "common knowledge" is that the thing we are sometimes concerned with is the consensus, whether or not that consensus is correct, or justified. Something like "common belief" would be a better term for this—so let me use that expression to cover those things that are believed, that a group largely believe each other to believe, and are aware of this mutual belief, etc. It is more than what is in fact believed by each of the group. (It may also be less, if there can be common knowledge or common belief in a group even if the occasional member does not believe it, or is not aware that it is commonly believed, or whatever). In determining what the "common belief" is, or what the public consensus is about a given topic, a platitudes analysis might be the best bet. Indeed, on one possible account of what platitudes are, the platitudes just are what is commonly believed (though facts about how they are to be weighted may go beyond information simply about what is commonly believed in a community).

A platitudes analysis might also be what is required when we want to find out what is believed by a group that we do not belong to: if I was trying to assemble the platitudes about wave-function collapse or shamanism or Russian literature I would not put a great deal of stock in what I believed. When we are trying to discover what the platitudes are among an expert community for some expression or concept, the result may well be a

---

<sup>8</sup> I don't mean to suggest that the people I talk to about philosophy are a sub-set of the people who agree with me—if that were the case, I wouldn't have philosophical conversations very often!

map of what are matters of widespread, relatively uncontroversial belief among that group. This information can be useful and interesting, even when it goes beyond telling us conceptual truths about the expression or concept in question. I claim that it can even be useful for advancing our knowledge or understanding of the world, and not merely the expression or concept—our understanding of wave-function collapse, shamanism or Russian literature, for example.

The characterisation of platitude-collection in the previous few paragraphs goes better with views of platitudes that are towards the “kitchen sink” end of the continuum discussed in the early sections of this paper. Presumably it is harder to read off which platitudes, if any, capture the unrevisable core than to tell what people, even experts, happen to believe, though perhaps we could at least suppose that it is a necessary condition for a claim being part of the unrevisable core of claims employing a concept that all the experts agree it is true. To the extent that part, or all, of assembling the platitudes is discovering a “folk theory” about a topic (or an “expert theory”, I suppose, if I am deferring to experts in a field), my remarks above should apply.

Self discovery about what I believe is all well and good—but how is it relevant to what is in fact the case? Suppose I do discover that I believe *X*, *Y* and *Z* about free will, or causation, or time. What sort of help is that in working out what free will, causation or time really are? (Setting aside the sort of benefits I mentioned in the previous section.) One way it helps is instrumental, or “pragmatic”—it can help to keep track of what you yourself think, while trying to improve your evidence about a given question. People who do not seem to realise exactly what they believe might get more easily muddled, or have a harder time in checking whether belief transitions they make during an investigation are justified. And being able to check one's belief transitions is important for most of us if we are going to avoid going wrong. Finding out what we believe, or what is “common belief”, or what the experts believe can be important as well. To the extent that metaphysics or anything else is a communal inquiry, one thing a lot of us as individuals will be wanting to do is to keep an eye on what the community of investigation thinks—the better to critique it, or extend it, or provide arguments to others

that proceed from premises they are willing to grant. Having good information about this sort of thing is like having time and funding—it isn't itself evidence of a view, but it is important to coming up with views that are closer to the truth. (We hope!)

I think it can also make an evidential difference what beliefs I have. Many writers have endorsed the view that we have some reason to continue to hold the beliefs we currently have, *because we currently hold them*, rather than change to a set of beliefs otherwise equally supported by the evidence. This is often called "epistemic conservatism". I should note that when I say it makes an evidential difference, I mean this in the sense that it makes for something epistemically better, *ceteris paribus*, rather than being of merely pragmatic use. (Or something close to *ceteris paribus* epistemic betterness, anyway—there is a useful debate to be had about how the evidential theoretical virtues relate to other epistemic evaluations.) Some adopt the usage where there are two sorts of ways a theory can be epistemically or methodologically good: one way is in relation to evidence, and the other is by having structural or theoretical features such as simplicity, conservativeness, etc. These people may well contrast “evidential” value with the sort of epistemic value I have in mind. My usage clearly differs from theirs, but so far as I can tell this difference is merely terminological.

Epistemic conservatism is common, though it is more commonly asserted to be good method (or presupposed to be good method) than argued for. Quine's principle of "minimal mutilation"<sup>9</sup> is in part a principle of theoretical conservatism—if new evidence requires me to shift the state of my beliefs, why ought I to make the minimum changes needed to restore coherence? Something like theoretical conservatism is also implicit in most Bayesian models of scientific inference. There are many admissible prior probabilities that one can have consistent with being rational, and they lead to different posterior probabilities when one processes evidence in the way Bayesians take to be rational. But even though this means that there are a range of posterior probabilities that are not ruled out by rationality, we are not rationally permitted to switch between these permissible posterior probabilities—we must stick to the ones we have, unless more

---

<sup>9</sup> Most famously outlined in Quine 1951, though he does not use the expression “minimal mutilation” there.

evidence comes in that requires us to update our probability distributions. In effect, this means that what priors someone in fact has (or had) constrains what credences it is rational for that person to have: what the person does believe, by itself, makes a difference to what they should believe in a conservative way—they are entitled to stay put, but not entitled to change, unless updating in response to evidence.<sup>10</sup> But theoretical conservatism is not simply a peculiarity of certain schools of philosophy of method—many "unaffiliated" (or less clearly affiliated) philosophers of science would agree with it.

At least one common methodology in philosophy in general seems to share in epistemic conservatism: the method of “reflective equilibrium”. While this method was also first articulated in the philosophy of science (see Goodman 1955), it was of course popularised under the name “reflective equilibrium”, particularly in ethics and politics, by John Rawls (Rawls 1971). “Reflective equilibrium” is a process of beginning with our particular judgments about a topic, together with the generalisations we are initially inclined to accept, and rules of adjudicating disputes, and by repeated application of each to the others coming to a final theory in which inconsistencies and tensions between our starting commitments are all smoothed out. Goodman, and following him Rawls, thought that justification would thereby be conferred on our final resting place, though a lively debate can be had about why, and whether a commitment to reflective equilibrium as the model of methodology brings with it a pressure towards some sort of constructivism or other anti-realism. (This debate is, I take it, a special case of the debate about whether coherentism in epistemology is in tension with realism in metaphysics.)

For our purposes, however, notice how reflective equilibrium suggests some form of epistemic conservatism. If one’s final point can be justified just because it comes from a starting point through a suitable process of harmonisation, and where one ends up depends heavily on where one began, then plausibly one can become justified in believing  $p$  at a later time largely due to the fact that one believed  $p$  at an earlier time.

---

<sup>10</sup> One broadly Bayesian view that is an exception is Bas van Fraassen’s voluntarism—see e.g. van Fraassen 1989 ch 7 for a discussion.

For example, it is quite plausible that two theorists could be rational and diligent practicers of reflective equilibrium, and one be justified in believing  $p$  but the other not be, where the difference is entirely due to the first including  $p$  in his starting commitments, while the second did not: the two could have starting points that are otherwise the same, and have done the same thing, *mutatis mutandis*, with their starting points. (Note well that this is a matter of epistemic justification, not merely some difference in appropriate pragmatic attitudes towards  $p$ .)

It is hard for me to be more specific about the link between reflective equilibrium and epistemic conservatism, partly because the details of what reflective equilibrium is and what justifies its application are not matters of widespread agreement. (Of course we could just select a specific account of reflective equilibrium and examine how much epistemic conservatism it implied: but that would tell us more about the particular choice of formulation than the general link.) One thing in particular that might be relevant is what the “method” of reflective equilibrium tells us about justification during the process: we are told we have justification at the end of the process, but is there anything in general to be said about the justification of our beliefs at any particular stage in the middle of the process? This question deserves more attention than it often gets, especially since in practice we are *always* in the middle of the process: nobody, not even Rawls, has completed the process of weighing up all our particular ethical judgments with all of our general principles and all of our methodological opinions about ethics. (Nor did Goodman ever weigh up *all* of our particular judgments about inferences, our general inferential commitments, etc. to give a final answer about what “inductive” inferences we should engage in.) Some tempting answers here, such as the claim we are justified in believing whatever has survived so far in the reflective process, at least absent evidence of serious error or incoherence, will give rise to fairly general principles saying that if we in fact believe  $p$ , we have some reason to continue to believe  $p$ , (perhaps with a few provisos to rule out justification in cases of particular epistemic depravity, or when a reasoner is sufficiently far from following the method of reflective equilibrium).

Epistemic conservatism deserves a defence in depth, especially given that many are

suspicious of it, and some philosophers have argued in detail against it (for a recent survey and attack on epistemic conservatism, see Vahid 2004). That would take at least a paper in itself, so for now I will be content to associate the view with its partners in crime: those already tempted by Quinean webs of belief, Bayesian updating mechanisms, or the practice of reflective equilibrium, should (nearly) all be tempted by epistemic conservatism in some form. A full discussion should also clarify exactly what sort of epistemic conservatism, if any, ought to be endorsed: in what circumstances does possessing a belief by itself oblige or permit an agent to continue to hold it? If it is a consideration to be weighed against others (as is surely plausible), how is this weighing to be done? These are all good questions, but they are all questions I want to sidestep here.

If epistemic conservatism is a virtue, then coming to realise what one believes is one way to come to realise what reasons one has for believing things as well: if the platitudes analysis reveals to me that I believe that properties are multiply located, then that shows me that I have some reason to think that they are.<sup>11</sup> To the extent that epistemic method is not individualistic, similar things will be true in the case of shared beliefs, "common belief", and the belief of experts: discovering that we believe something makes a difference to what we have reasons to believe. (Whether what we need to discover is what we each happen to believe in common, or something stronger like "common belief", is a question about collective reasons for belief I need not tackle here.) We must be a little careful here—if having the belief is enough to ensure there is a reason, then the platitudes analysis does not generate that reason—it is already there, all the platitude analysis would do in this case is *make us aware of* a pre-existing reason for holding the beliefs in question. So the platitudes analysis does not give us these kinds of reasons (except insofar as putting commonplaces together may lead us to clean up our beliefs), but it only gives us explicit access to the reasons (or their grounds). However, if the epistemic conservatives are right, and having a belief is itself something that gives us a reason to hold that belief, then beliefs about whether or not one believes *p* themselves

---

<sup>11</sup> That properties are multiply located is not in fact something I think a platitudes analysis would reveal is part of a folk theory of properties.

may give one an indirect reason for believing  $p$ . This is because taking yourself to believe  $p$  is good, though perhaps not indefeasible, evidence that you do in fact believe  $p$ . So the belief that I believe  $p$  is evidence that I believe  $p$ , so it is evidence that I have evidence that  $p$ . I am inclined to think that evidence of evidence is itself evidence (that is, evidence of evidence that  $p$  is itself evidence that  $p$ ). This may not be obvious (in methodology, what is?) but some examples might help make it plausible:

1. I have a hazy memory that I proved Conjecture X to my satisfaction last week, though I cannot presently recall the details.
2. A competent and reliable informant tells me that he has new evidence that Jones committed the murder, though he cannot tell me what that evidence is without breaking a confidence.
3. I have a visual impression of the shadow of a stick reaching the two-metre mark on a measuring tape. I happen to know the sun is at a 45 degree angle to where I am standing, so I take it that the stick is two metres long.

In each of these cases, I take it I have evidence of evidence of something—a memory of a deduction, testimony of some-evidence-or-other, and perceptual evidence of a measurement which itself is (in context) evidence for the height of the stick. In each of these cases I take it is quite plausible that the evidence of evidence for  $p$  is itself evidence: of Conjecture X, that Jones is a murderer, and that the stick is two metres long. If evidence of evidence were not itself evidence, then we would have to be very careful about many of our several-step chains of discovery.<sup>12</sup>

If coming to believe that I believe that  $p$  is then itself evidence for  $p$ , then the platitudes analysis can generate reasons to believe propositions about the world, rather than merely reveal that we have reasons we may not have noticed. I think this is one of the significant things going on when we engage in thought experiments and intuition-pumps whose conclusions are not just about our concepts, but about how the world is.

---

<sup>12</sup> Exercise for the reader: see if you can come up with cases where evidence of evidence that  $p$  is not itself evidence for  $p$ .

Some might think that if epistemic conservatism has the result that getting new evidence for the world is this easy, that is a *reductio* of epistemic conservatism. I do not intend to defend epistemic conservatism in depth here, but it is worth pointing out that this need not be very strong evidence at all: for all I have said here, it need not ever be enough to justify confident belief on its own. Even if it is not, it may serve as a valuable methodological starting place—if we are permitted to continue in our beliefs until we work out a better alternative, that may be preferable to having to engage in widespread suspension of judgment. A commitment to epistemic conservatism, even with the additional commitment that evidence about our own beliefs is itself evidence of the truth about the subject matter of those beliefs, need not bring with it the suggestion that this is a particularly fruitful way to proceed in practice—I suspect it will be of most use at early stages of inquiry, before we engage sufficiently with a question to be able to do better than rest content with our starting place. It may also be helpful to point out the analogy with testimony—the discovery that others believe that *p*, e.g. by their testifying that *p*, is often some evidence that *p*, even if that evidence is eminently defeasible and may count in favour of *p* only because of contingent things we assume about the reliability of other people. And I suspect this reliance on others is fairly fundamental—it often, of necessity, provides the starting point of our opinions about the world rather than its being the case that we only take others’ opinions into account after carefully examining their *bona fides* and independently established reliability. Is it so terrible to extend to ourselves the same courtesy that we extend to the woman or man on the street?

Finding out what we think may serve one more related goal, but one that is probably less valuable. Contemporary philosophy is sometimes conducted in an adversarial manner, where we see the job of a view’s proponent to make the most persuasive case possible. Other times we have some tendency to treat it as a rhetorical exercise, where the job is to convince our readers or interlocutors of a view (or more feasibly, to get them to find the view more plausible than they did before). In this environment, we may end up just trying to convince people of our views, and lose sight to some extent of the question of what we have good reason to believe. So sometimes, I imagine, discovering what someone

believes, or what we collectively believe, is done only to make someone consciously convinced of something: we may on occasion use an intuition pump or extract a consequence from folk theory, not just because we are aiming at getting a better theory, but simply to convince an interlocutor (or reassure ourselves).

I have suggested that when this is going on, often we are contributing to our theory's value—either evidentially, or perhaps by improving some pragmatic virtue. But this use is primarily *ad hominem*: we are merely trying to find beliefs that the party-to-be-convinced holds, or can be convinced to hold, that can be used as premises—and trying to do this merely because it is a way to get the person in question to agree with us. This is dialectically useful, and no doubt most of us stoop to this at some stage or another—but simply persuading someone that one's view is correct is not necessarily very valuable unless one is bringing out *reasons* to believe the view being supported. (Not valuable if we are trying to find the truth, that is—there are other motives for trying to get people to agree with you, of course.)

I expect the opponents of epistemic conservatism will be more inclined to think that this is all that is going on when we are articulating views we may be holding implicitly, or make appeals to others by bringing out features of "common belief" or features of the views they implicitly hold. I think we should be on our guard, when presented with arguments from premises we accept, to make sure we have good enough reasons to believe the premises before we become too convinced of the conclusions. I can afford to be a little indulgent of my beliefs, given that I take epistemic conservatism to be a theoretical virtue, and an evidential theoretical virtue, at that: but those who do not had best be especially careful. Nevertheless, given that *in fact* we can secure agreement with another by convincing that person that they believed the proposition in dispute all along, no doubt the role that platitudes analysis can play in doing this is relevant to the philosophical enterprise as it is in fact practiced.

### *The Third Way*

Discovering that one already has an opinion may in general be of some help in investigation—either because it is evidentiary, or for some more pragmatic reason. But it is particularly useful when one has some reason to suppose that people are good at coming to know the sort of thing the opinion is about. This applies to cases of conceptual analysis, presumably—one reason we think that analysis of concepts or expressions yields knowledge is that we have some useful access to our own concepts, or (perhaps more dubiously) to the meanings of expressions in public languages that we speak. If I find myself inclined to think that it is true in virtue of the meaning of the word "brother" that anything that is a brother is a sibling, then part of why I take myself to be justified is that I think I know what "brother" means in English. The claims that we know a lot about the concepts we use and the meanings of the words we use are both controversial claims, and some theorists have denied one or both. But at least the traditional picture of conceptual analysis or linguistic analysis seemed to presuppose that we did reliably have such knowledge (though it was of course defeasible and perhaps limited in other ways).

We may have reason to suppose we are reliable about questions other than questions of meaning, of course. We may also on occasion have reason to think our pre-theoretic judgments are to be trusted—or at least those judgments made before explicit, articulated theorising are to be trusted, since we may have good reason to think our supposedly "pre-theoretic" judgments are embedded in an implicit theory. Some questions in metaphysics might be thought to be like this—even if we find it very hard to articulate views or to know how to sort the true from the false ones, we might have reason to suspect we already know something about the subject that we are having trouble articulating, rather than thinking that we do not know what is going on at all. What it is to be an agent in the world, for example, is at least partly a metaphysical puzzle, and metaphysicians argue over whether we need to expand our metaphysical commitments to provide a satisfactory theory of agency. (Are an agent's decisions after all some complex of physical and biological events, or something constructed from them in an innocent way? Or is there something fundamental and metaphysically special going on when an agent makes a choice that is absent from other phenomena?) Debates about freedom of the will, the metaphysics of mind and intentionality, the metaphysics of morals and normativity, etc.

often touch upon this difficult-to-answer question. We might expect that we implicitly know quite a bit about agency—after all we are agents, we are good at predicting and understanding the behaviour of other agents (each other), and there is good reason to suppose that we even underwent evolutionary pressure to reliably track some things about agency—those without the self-awareness to monitor their decisions, or the social awareness to understand and anticipate others, will often have done less well. We might therefore take more seriously the deliverances of “common sense” about agency than we would or should take the deliverances of “common sense” about what happens to objects as they approach the speed of light or when they are one nanometer across.

Another example concerns abstract ontology. Before we do much explicit theorising, we have a lot of beliefs about modal matters—we have beliefs about what is possible, we have habits of ascribing dispositions to things, and we are able to make conditional claims, and take them to be justified or unjustified in particular cases, well in advance of being able to articulate general principles about them or to answer some of the questions we would like general theories of modality to answer. Examples might include our conviction that if something necessarily occurs than it possibly occurs, or that discovering that something is in fact the case shows that it is possibly the case. On the other hand, our ordinary practices and ordinary ways of talking seem to have very little to tell us about what large cardinal claims in set theory are true (i.e. claims about how different very large infinite sizes of sets are related). A case can be made that we should take our platitudinous attitudes towards modality more seriously as evidence than we should take our initial opinions about large cardinal axioms in set theory.

So the third way the platitudes analysis might help us is when it reveals to us a theory that we have some independent reason to believe is likely to be reliable. We may in fact be in a position to take a theory to be justified or reliable even before we can properly say why it is reliable or why it gives us warrant—I presume we can trust the evidence of our eyes before we have a good story about the epistemology of perception, and we can often rely on the testimony of others before we know the details of how or why they come to be authoritative reporters. Something similar may be true when we come to answer

questions in the metaphysics of agency or time or causation—I am inclined to think our implicit grasp of causation is better than our explicit ability to answer questions about causation, which is one of the reasons why I think intuition pumps and assembling platitudes about causation is a worthwhile exercise. (Though I am inclined to think that we will eventually be able to do better than present cleaned-up folk wisdom about what it is for one thing to cause another.)

***Are any of these 3 things what people are actually doing?***

It is comparatively rare for people to explicitly take themselves to be employing a platitude account when doing philosophical analysis. Many people are not even clear what they take themselves to be doing when they are writing about metaphysics—they may call something they put forward an "analysis", or they may rely on an "intuition pump", or they may offer a counterexample to a theory without being explicit about what their justification for taking the counterexample to be correct is. Whether in such cases people take themselves to be doing conceptual analysis, or applying the results of another investigation (e.g. a result from the natural sciences), or doing something else, is not always clear.

While I would sometimes like people to be clearer, I do not even want to claim that people should always be. Not every paper is a methodology paper, and if we insist that every time someone gives us an argument they also give us the methodology behind that argument, we may end up in an unpleasant regress (e.g. if we expect people to justify their methodologies as well, and justify that justification, etc.) Or we may just end up with less satisfying work on the original problem—methodological excursus is sometimes distracting, and we are usually worse at saying why X is a reason for Y than telling whether X is in fact a reason for Y. So sometimes at least it is fine for a metaphysician (or philosopher, or other theorist) to just get on with doing what they are doing, and we can work out what their methodology was or should have been after the fact (if we bother to explicitly work it out at all).

So my guess is that some metaphysicians are "just getting on with it": thought experiments, counterexamples, and plausible commonplaces are employed, and whether they work through being conceptual analysis, or a deployment of folk theory, or something else.<sup>13</sup> I offer these three ways platitudes can help not as a story about what metaphysicians think they are doing; nor even, necessarily, as a story about what they are doing, though keeping them in mind while answering that question may be helpful. Instead, they are offered as justifications for practices that many of us are engaged in.

The alert reader will have noticed that my remarks about metaphysics and the Canberra Plan have wider application than to the activity of Canberra planners and fellow travelers. Anyone interested in fundamental metaphysics and interested in methodology should be interested in the question of how armchair methods can help us. Someone can disagree with various parts of the Canberra plan project, or alternatively be relatively uninterested in where our everyday notions fit into the world as described in more fundamental vocabulary, and still see the merit in the three ways discussed that the armchair can help in fundamental metaphysics. Indeed, people who are not even particularly interested in metaphysics might find this discussion useful, since virtually no matter what inquiry one is engaged in, the question of how "relatively a priori" investigation can help may be interesting. Perhaps these sorts of contributions in some areas are overshadowed by other methods: but investigating the world can be hard, and it would be a shame to neglect any source of epistemic justification we have available.<sup>14</sup>

*Daniel Nolan*

*Departments of Philosophy*

---

<sup>13</sup> Another thing that I think one can do from the armchair is apply theoretical virtues such as simplicity, unification, and perhaps some others: many of these can be subsumed under a category of considerations that make for a better explanation, in the "inference to the best explanation" sense. Some claim that these virtues are only pragmatic: I am inclined to think many are evidential, but in either case as long as theories can be improved by their application, applying these standards (e.g. by arguing some alternative is simpler or more unified than another) can be worthwhile—and applying these standards can often be usefully done without additional data collection.

<sup>14</sup> Thanks to Chris Daly, Alan Hájek, Katherine Hawley, Ishani Maitra and audiences at the 2003 Metaphysics in Science Workshop in Athens, the 2003 AAP conference in Adelaide, and at the Universities of Glasgow and Leeds, La Trobe University and UC Davis, for comments and discussion. Especial thanks to the staff and students in the Philosophy Program at ANU from 1994-1996, for Planning.

*University of St Andrews*  
*St Andrews, Fife, Scotland*  
*United Kingdom*  
*dpn@st-andrews.ac.uk*

### **References**

- Braddon-Mitchell, D. and Jackson, F. C. 1996. *Philosophy of Mind and Cognition*. Blackwell, Oxford.
- Burge, T. 1979. "Individualism and the Mental" in French, Uehling, and Wettstein (eds.) *Midwest Studies in Philosophy IV*, University of Minnesota Press, Minneapolis. pp 73-121.
- Carnap, R. 1963. "Replies and Expositions" in Schilpp, P.A. (ed). 1963. *The Philosophy of Rudolf Carnap*. Library of Living Philosophers, La Salle, Illinois. pp 859-1013
- Goodman. N. 1955. *Fact, Fiction and Forecast*. Harvard University Press, Cambridge MA.
- Jackson, F.C. 1994. "Armchair Metaphysics" in Michael M. and O'Leary-Hawthorne, J. (eds). 1994. *Philosophy in Mind*. Kluwer, Dordrecht, pp 23-42
- Jackson, F.C. 1998a. *From Metaphysics to Ethics*. Oxford University Press, Oxford.
- Kripke, S. *Naming and Necessity*. Basil Blackwell, Oxford.
- Jackson, F. C. 1998b. "Reference and Description Revisited". *Philosophical Perspectives* 12: 201-218.
- Lewis, D.K. 1970. "How to Define Theoretical Terms", *Journal of Philosophy* 67: 427-46. Page references are to the reprint in Lewis 1983.
- Lewis, D.K. 1972. "Psychophysical and Theoretical Identifications". *Australasian Journal of Philosophy* 50: 249-58. Page references are to the reprint in Lewis 1999.
- Lewis, D.K. 1983. *Philosophical Papers, Volume I*. 1983. Oxford University Press, Oxford.
- Lewis, D.K. 1989. "Dispositional Theories of Value". *Proceedings of the Aristotelian Society, Supplementary Volume* 63: 113-37

- Lewis, D.K. 1999. *Papers in Metaphysics and Epistemology*. Cambridge University Press, Cambridge.
- Lewis, D.K. 2000. "Causation as Influence". *Journal of Philosophy* 97: 182-197
- Nolan, D. 2005. *David Lewis*. Acumen, Chesham.
- Price, H. 2004. "Naturalism Without Representationalism" in Macarthur, D. and de Caro, M. (eds) 2004. *Naturalism in Question*. Harvard University Press, Cambridge MA. pp 71-88
- Price, H. and O'Leary-Hawthorne, J. 1996. "How to Stand Up For Non-Cognitivists". *Australasian Journal of Philosophy* 75: 274-92
- Putnam, H. 1975. "The Meaning of 'Meaning'" in Gunderson, K (ed) *Language, Mind and Knowledge: Minnesota Studies in the Philosophy of Science* 7, University of Minnesota Press, Minneapolis. pp 131-196
- Quine, W.V. 1951. "Two Dogmas of Empiricism". *Philosophical Review* 60.1: 20-43
- Rawls, J. 1971. *A Theory of Justice*. Harvard University Press, Cambridge MA.
- Smith, M. 1994. *The Moral Problem*. Blackwell, Oxford.
- Vahid, H. 2004. "Varieties of Epistemic Conservatism". *Synthese* 141: 97-122
- van Fraassen, B.C. 1989. *Laws and Symmetry*. Clarendon, Oxford.
- Wright, C. 1998. "Truth: A Traditional Debate Revisited". *Canadian Journal of Philosophy*, supplementary vol. 24: 31-74