A Phenomenal, Dispositional Account of Belief

Eric Schwitzgebel
Department of Philosophy
University of California
Riverside, CA 92521-0201
eschwitz@citrus.ucr.edu
http://www.ucr.edu/philosophy/schwitz.html

Decemer 4, 2000

A Phenomenal, Dispositional Account of Belief

Most philosophers of mind these days think that to believe something is to have a representation of some sort in one's mind.[1] A <u>prima facie</u> difficulty with this view is that representations seem to be discrete entities – either fully present in the mind or wholly absent – while belief is a continuous phenomenon. Not only does believing come in degrees of certainty (a widely-acknowledged fact but one that introduces substantial complications for a representational account of belief if pursued in detail[2]), but also multifarious other cases exist in which people are neither quite accurately describable as believing a certain proposition, nor quite accurately describable as failing to believe it, for reasons that go beyond uncertainty, such as in cases of self-deception, gradual learning, forgetfulness, and when there is ignorance of related facts. If the reader has any trouble imagining such cases, a number will be described in this paper.

I think that the preceding difficulty for the representational view of belief is more than just <u>prima facie</u>, though I will not argue the point here.[3] What I propose to do instead is to suggest a way of thinking about belief harmonious with the fuzzy, continuous nature of the phenomenon, an account that provides the tools usefully to describe both central cases of believing and peripheral cases in which the subject stands somewhere between believing and failing to believe. In offering such an account, I take myself to be doing something like what Carnap (1950) called "explication"; I am presenting a way to think about belief that may not quite match our ordinary understanding (which is a bit confused) but comes close and, hopefully, has the compensating virtues of precision, clarity, and utility for philosophical and scientific purposes.

I call the account of belief I am about to offer a <u>phenomenal, dispositional</u> account. I call it a <u>dispositional</u> account because it treats believing as nothing more or less than being disposed to do and experience certain kinds of things. I call it a <u>phenomenal</u>

account because, unlike dispositional accounts as typically conceived, it gives a central role to first-person, subjective experience, or "phenomenology." I will begin with a statement of the account. I will then discuss "in-between" cases of believing in some detail. I will conclude with a discussion of the relations between the proposed account of belief and several other positions one might take regarding belief.

1. The Account.

To begin, we will need the concept of a <u>dispositional stereotype.</u> By a <u>stereotype,</u> I mean a cluster of properties we are apt to associate with a thing, a class of things, or a property.[4] Adapting an example from Putnam (1975), stereotypical properties of tigers include (among other things) their being striped and their being four-legged. A white or three-legged creature may still be a tiger, but it cannot be a stereotypical one. Stereotypes may be accurate to different degrees. An accurate stereotype is one in which most of the stereotypical properties are had by most of the objects to which the stereotype applies; a less accurate stereotype is one in which the stereotypical properties are instantiated in fewer of the objects to which the stereotype applies. Some elements of a stereotype may be broadly inaccurate – for example, if one of the stereotypical properties of tigers were that they lived in African jungles.

    <u>Dispositions</u> can be characterized by means of conditional statements of the form: If condition C holds, then object O will (or is likely to) enter (or remain in) state S. O's entering S we may call the <u>manifestation</u> of the disposition, C we may call <u>condition of manifestation</u> of the disposition, and the event of C's obtaining we may call the <u>trigger.</u> Exactly what the connection is between O's having the dispositional property to enter state S in condition C and the truth of the conditional statement associated with that disposition is a matter of some debate, into the complexities of which I will not enter here as the account does not depend on one particular way or other of resolving the matter.[5]

As a rule of thumb, however, we may suppose that O has the disposition in question just in case the corresponding conditional statement is true. Thus, for example, salt has the dispositional property of being soluble in water because it is apt to dissolve (the manifestation) upon being placed in water (the trigger). Mirrors are disposed to reflect light because when light shines on them (the trigger), it reflects back (the manifestation). Carlos is disposed today to get angry when his car doesn't start because if his car doesn't start today, he is likely to get angry.

A <u>dispositional stereotype</u> is a stereotype whose elements are dispositional properties. An example is the stereotype for <u>being hot-tempered</u> This stereotype will include the disposition to respond angrily to minor provocations, the disposition to be slow in cooling off after a fight, the disposition to feel and express frustration quickly when one's will is thwarted, and so forth. Personality traits, such as being reliable, affable, or tenacious, are arguably all characterizable by means of such dispositional stereotypes. To have these personality traits is really nothing more than to match these stereotypes. The core claim of this paper is that belief can be characterized in much the same way.

Consider a favorite belief of philosophers: the belief that there is beer in the fridge. Some of the dispositions associated with this belief include: the disposition to say, in appropriate circumstances, sentences like 'There's beer in my fridge;' the disposition to look in the fridge if one wants a beer; a readiness to offer beer to a thirsty guest; the disposition to utter silently to oneself, in appropriate contexts, 'There's beer in my fridge'; an aptness to feel surprise should one go to the fridge and find no beer; the disposition to draw conclusions entailed by the proposition that there is beer in the fridge (e.g., that there is something in the fridge, that there is beer in the house); and so forth.

Since a list of dispositions of this sort may be continued indefinitely, the dispositions stereotypically associated with a belief cannot all, or even commonly, be associated with it explicitly, as the result of some conscious process of linking each belief

with each disposition in its stereotype. Think of the dispositional stereotype for the belief that P, rather, as consisting of the cluster of dispositions that we are apt to associate with the belief that P. Dispositional personality traits may similarly not be amenable to full explicit specification: How exactly is a smart person disposed to think and act? We can start a list of stereotypical dispositions, but as in the belief case, it seems impossible to finish it without circularity. Furthermore, due to the generativity of language, the number of possible beliefs, and so the number of belief stereotypes, is indefinitely large. Even if no one has ever considered the possibility of believing that Ed's bowling ball is stuffed with custard, commonsense psychology equips us amply enough that we are still apt to associate certain dispositions with that belief. Therefore, I would like to distance the technical concept of 'stereotype' employed here from any elements of the ordinary notion of a stereotype that preclude the possibility of stereotypes that have never been used.

A further complication arises from the fact that not every person will be apt to associate exactly the same dispositional properties with any given belief – no more than everyone would agree about the stereotypical properties of a chair. For this reason, it is important to talk about a "cluster" of stereotypical properties: Some properties will be central and widely agreed upon, others will be marginal. Because a preponderance of the most central properties in the stereotype for a chair can be widely agreed upon, we can agree to a substantial extent about what are and are not typical chairs. A similar general but imperfect concord is available in the case of belief. For simplicity's sake, then, I will generally speak as though there were one stereotype for every belief, though strictly speaking this is not true.[6]

The dispositional properties belonging to belief stereotypes fall into three main categories. The most obvious, perhaps, are behavioral dispositions, the manifestations of which are verbal and nonverbal behavior, such as, in the present case, the disposition to say that there is beer in the fridge (in appropriate circumstances) and the disposition to go to the fridge (if one wants a beer). Equally important, though rarely invoked in

dispositional accounts of any sort, are what may be called <u>phenomenal</u> dispositions, dispositions to have certain sorts of subjective, phenomenal experiences. The disposition to say silently to oneself, "there's beer in my fridge," and the disposition to feel surprise should one open the fridge and find no beer are phenomenal dispositions stereotypical of the belief that there is beer in the fridge. Finally, there are dispositions to enter mental states that are not wholly characterizable phenomenally, such as dispositions to draw conclusions entailed by the belief in question or to acquire new desires or habits consonant with the belief. Call these <u>cognitive</u> dispositions. Some important stereotypical dispositions may be hybrids of two or more of the preceding types, irreducible into components belonging wholly to a single category, such as the disposition to search anxiously through the fridge if the beer seems unexpectedly to have run out.[7]

On my view, a person who possesses all the dispositions in the stereotype for (for example) believing that "There is beer in my fridge" can always accurately be described as believing that there is beer in her fridge. A person who possesses none of the relevant dispositions can never accurately be so described. And, of course, between these two extremes is a wide range of cases in which the subject has some but not all the dispositions in the stereotype. Roughly speaking, the greater the proportion of stereotypical dispositions a person possesses, and the more central those dispositions are to the stereotype, the more appropriate it is to describe her as possessing the belief in question. An additional element of vagueness is introduced if one accepts that having a disposition is not itself always a simple yes-or-no matter.

No one disposition is either necessary or sufficient for the possession of any belief. It may occur to some that the disposition to feel assent to an internal utterance of P (or to think silently to oneself, "P") comes close to being a sufficient condition for believing that P; nevertheless, we must allow that people sometimes feel assent to utterances the contents of which it is not wholly accurate to describe them as believing, for example,

when they don't really understand what the utterance means or when they are self-deceived.

To believe that P, on the view I am proposing, is nothing more than to match to an appropriate degree and in appropriate respects the dispositional stereotype for believing that P. What respects and degrees of match are to account as "appropriate" will vary contextually and so must be left as a matter of judgment. This vagueness and context-dependency does not undermine the value of belief ascription, but rather makes it flexible and responsive to our needs as belief ascribers. Similar vagueness and context-dependency can be found in the ascription of character traits, providing them with a similar flexible utility. The numerous examples in this paper will, I hope, help to support the view that talk about belief can be vague and flexible in this way and still quite useful – more useful in fact than an approach that rigidly insists on determinate yes-or-no answers to all questions about what people believe. I will describe a number of in-between cases of believing in section four, but first I will discuss an issue pertinent to handling cases of deviation from the stereotypes and compare the account with earlier dispositional accounts of mental states.

2. **Ceteris Paribus** Clauses and Excusing Conditions.

A substantial complication for this account, crucial to the assessment of adherence to the stereotypes, arises from the fact that the stereotypical dispositions to which I appeal hold only ceteris paribus or "all else being equal." Joe might believe there is beer in his fridge, but if he is particularly stingy with his beer, he may not be ready to offer beer to a thirsty guest or even to admit he possesses beer at all. We wouldn't want to say that Joe's reluctance to divulge the presence of beer in his fridge makes it any less accurate to describe him as believing it to be there. Behavioral dispositions, or dispositions with a behavioral element, seem particularly defeasible in this manner. It is no challenge to

concoct cases in which the explicit conditions of manifestation of a stereotypical disposition are satisfied, yet despite the clear presence of the pertinent belief, the behavior does not manifest. It is usually sufficient, though crude, to imagine a well-armed sociopath insisting that the behavior not take place. Purely phenomenal dispositions may be less defeasible in this way. It is harder (but not impossible, I suppose) to imagine cases in which it seems right to say that Joe fully and completely believes that there is beer in his fridge, yet he feels no surprise upon opening his fridge and noticing a lack of beer.

For the account I am offering to be plausible, it must acknowledge that the dispositions in belief stereotypes hold only <u>ceteris paribus</u>. When the situation is nonstandard in certain ways, the stereotypical dispositions cannot be expected to manifest. One way of thinking about this matter is to view the dispositional characterizations as loaded with tacit "if" clauses. Not literally <u>all</u> else must be equal – but certain conditions must hold. Joe, for example, is disposed to assent to utterances meaning that there is beer in his fridge <u>if</u> he hears the utterance, <u>if</u> he has not decided to lie about or evade the matter, <u>if</u> he understands the language in which the utterance takes place, <u>if</u> he has the physical capacity to indicate assent, and so forth. In assessing deviation from a stereotype, the evaluation of such potentially excusing conditions is crucial.

Many scientific and ordinary generalizations are similarly <u>ceteris paribus</u> defeasible. Traffic on L.A. freeways will slow down around five o'clock on weekdays, <u>if</u> it is not a holiday, <u>if</u> L.A. is not being attacked by a foreign power, <u>if</u> a freak snowstorm hasn't made the freeways impassable, etc. Rivers erode their outside bank at a bend <u>if</u> the river is not frozen, <u>if</u> the bank is made of an erodible material, <u>if</u> there isn't a powerful fan in place preventing the river from touching the outside bank, etc. The <u>ceteris paribus</u> nature of such generalizations does not in these cases, nor I think in the dispositional case (a

dispositional claim is, after all, plausibly a type of modal generalization), hinder their productive use.[8]

I will treat it as an open metaphysical question whether dispositional manifestations must always occur when all the conditions of manifestation are satisfied. If so, then dispositions must often have an indefinitely large number of tacit conditions. The conditions of manifestation of many dispositions must, if completely fleshed out, be indefinitely long conjunctions (unless we cut the matter short by adding something like "and nothing prevents it" to the conditions of manifestation – but some would consider that to be cheating). Alternatively, one may wish to regard only a few conditions as true conditions of manifestation of any given disposition, if one is not averse to the idea that dispositions do not always manifest themselves when their conditions of manifestation are satisfied (see Martin and Heil 1998). With either approach, one may choose to be stingy or generous about how many of the conditions of manifestation of a disposition are to be made explicit and how many to be left tacit. My account is flexible on these points, though my inclination is to keep the dispositions relatively simple and folksy rather than to load them with lots of tacit conditions.[9]

A person may then be <u>excused</u> from a dispositional manifestation – i.e., not seen as deviating objectionably from the dispositional stereotype – if one of the tacit conditions of manifestation is not met, or (metaphysics permitting) if the disposition is not manifested for some other reason consistent with the possession of the disposition. Certain types of conditions are regularly regarded as excusers in this sense, such as physical incapacity or the presence of a desire or situation that makes a particular manifestation prudentially inadvisable. If Joe's speaking would set off a hair-trigger bomb, it does not count against the accuracy of describing him as believing that there is beer in the fridge that he will not tell us so. Other conditions may be somewhat less excusing and are apt to propel us into vagueness about what the subject believes, such as ignorance about related topics (e.g., Joe believes that Budweiser is not a type of beer),

distraction by other cognitive demands, or apparent failure to reason correctly. If Joe knows that his fridge is packed with Budweiser but does not think Budweiser a type of beer, does Joe believe that there is beer in his fridge? Those fond of the de re/de dicto distinction might remark that Joe seems to believe (de re) of a certain type of beer that it is in his fridge, but not to believe (de dicto) that there is beer in his fridge. This is only one way – and a questionable one (see Stich 1983 and Dennett 1987) – of trying to get a handle on intuitions that pull us in different directions in such cases. As the paper develops, I hope it will become clear how my account gives us another way of thinking about such vague, in-between cases of believing.

One may want to find a single, unifying principle that can guide us in distinguishing the lack of a disposition (and thus potential deviation from a stereotype) from excused non-manifestations of that disposition. This is essentially a demand for a principle unifying all the tacit conditions of or ceteris paribus excusers from dispositional manifestation. I think the prospects for finding such a principle are slender, but a brief detour to look at the question is nonetheless instructive.

Let's begin with examples of dispositions pertinent to Joe's belief that there is beer in his fridge. Normally, if the disposition in question would have manifested itself but for the presence of some hindrance outside of Joe, we are ready to grant excuses. If Joe doesn't offer beer to a guest only because of the bomb, we don't count that fact against the accuracy of describing him as having the belief. A general shutdown of the mind also seems to be excusing: We don't blame Joe for not offering the beer if he has momentarily blacked out. On the other hand, if Joe denies having beer in his fridge when a guest requests some, and we cannot attribute his denial to any external cause, nor to an intention to lie, nor to a misunderstanding of the question, his deviation may be unexcused, and it is natural to conclude that he doesn't realize that there is beer in the fridge. Joe's deviation from the stereotype in this case suggests, in a way the deviations

discussed at the beginning of this paragraph do not, an enduring likelihood of his deviating from diverse other aspects of the dispositional stereotype as well.

This last point may seem to hold some promise for the construction of a general principle differentiating excused non-manifestations from unexcused deviations. In cases of linguistic misunderstanding, or of deliberate concealment, or of yielding to external pressures, failure to manifest a stereotypical disposition does not seem to be symptomatic of a broad and enduring behavioral, phenomenal, and cognitive nonconformity to the stereotype. For example, Joe might well be uttering silently to himself, in any of these cases, "There is beer in my fridge." We have no reason, in such cases, to expect a <u>general</u> non-adherence to the stereotype; the deviation is naturally confined to a particular range of circumstances. If the gunman were to walk away, if his guest were to start speaking English, if Joe did not feel his precious beer threatened by the presence of a thirsty visitor, we would again see general conformity to the stereotype. One can treat cases of general mental or physical shutdown similarly, if one thinks of these conditions as particular, narrow circumstances. Perhaps, then, some idea of containment of the deviation could be drafted to serve as a general principle for identifying excusing conditions.

The question then arises, however, whether in putting forward such a principle we have added anything of substance to the account. Scientific and everyday generalizations are shown false by deviations that undermine our reasons for thinking the generalizations to be widely, approximately, or at least in "ideal" circumstances, right; we introduce <u>ceteris paribus</u> excuses in just those cases where we feel that a deviation from the generalization does not affect its overall applicability. Introducing a rule, then, that says <u>ceteris paribus</u> excuses are to be admitted exactly when a deviation does not threaten the basic accuracy of the generalization is simply to state what is implicit in the admission of <u>ceteris paribus</u> defeasibility from the beginning.

Clarifying this point helps us to see two factors that come together in assessing deviations as potentially excused. The first factor is an empirical assessment of the

likelihood of the generalization's broadly falling apart given that the deviation has taken place – in other words, how broad or narrow is the range of circumstances in which we can expect the generalization not to hold? The second is an assessment of the importance of deviation in the circumstances in which deviation can be expected. Can one afford a certain amount of looseness in the generalization because the cases are marginal or covered by other generalizations, or will one want to insist on a stricter regularity? No set of explicit rules seems to be able to guide us as well in making such assessments as does a well-practiced, intuitive grasp of the topic in question. This lack of explicitly specifiable rules for separating excused from unexcused deviations from a generalization suffuses even the most robust scientific theories (see, e.g., Cartwright 1983). Philosophers of science have learned to resist the temptation to try spelling out in full detail the ceteris paribus conditions for substantive, specific, scientific generalizations.

In the case of belief, then, the manifestation of stereotypical dispositions depends on the satisfaction of both a range of explicit conditions and a range of tacit, ceteris paribus conditions that may not be fully specifiable. When a subject does not, or would not, manifest a stereotypical disposition despite the satisfaction of all explicit conditions of manifestation, we may say that the subject deviates from the stereotype. A deviation is excused just in case an implicit condition of manifestation is not met or the manifestation is prevented by some circumstance consistent with the possession of the disposition – in other words, if the ceteris paribus clause is sprung. Although there can be no firm rule to distinguish excused from unexcused deviations, we have an intuitive sense that allows us to sort deviations along a spectrum from the clearly excused (such as deviations due to physical incapacity) to the clearly unexcused, with vague cases in the middle.[10]

If a deviation from a stereotype for a belief is excused, it counts not at all against the accuracy of describing the person as having that belief. If it is not fully excused, the question of whether it will count as an important deviation – one that makes us hesitate to ascribe the belief or makes the belief ascription less apt than it could be – will generally

depend on the context of the belief ascription. Suppose, for example, that a child studying for a test reads, "The Pilgrims landed at Plymouth Rock in 1620," and remembers this fact. She is a bit confused about what Pilgrims are, though: She is unsure whether they were religious refugees or warriors or American natives. Now does she believe that the Pilgrims landed at Plymouth Rock in 1620? She deviates from the stereotype in some respects: She will not conclude that Europeans landed at Plymouth Rock in 1620; and when she imagines the event, she may bring some inappropriate images to mind.   In some contexts – e.g., if we are talking about her likely performance on a history dates quiz – we might be inclined to describe her as believing this fact about the Pilgrims; in other contexts we would not. I am not here saying that the mental state of the child varies with context. Rather, given that the child deviates from the stereotype in some respects but not in others, how best to <u>describe</u> her mental state will depend on the practical demands of the moment.

This context-dependence is an important feature of the proposed account. Different dispositional properties will, in different contexts, be more or less crucial to decisions about whether to ascribe a particular belief or not, and in intermediate cases failure to attend to the context of ascription can result in differing assessments of the appropriateness of a belief ascription. Such inattention to context may be partly responsible for the wavering and disagreement about how best to describe the kinds of in-between cases of believing that philosophers sometimes find puzzling.

3. Differences from Traditional Dispositional Accounts and a Thought on Ryle.

Dispositional accounts of mental states are not, of course, new. Ryle (1949) helped begin a trend toward regarding much of mental life as fundamentally dispositional or dispositionally specifiable – or, not so differently, as functionally specifiable.[11] (One might think of a dispositionally specifiable state as a state of an object, e.g., a brain, apt to

bring about specified effects under specified conditions, and of a functionally specifiable state as a state of an object apt to bring about specified effects under specified conditions and to be produced by specified causes.) Others (e.g., Marcus 1990, Searle 1992, Audi 1994) have endorsed dispositional accounts of belief, or specifically of unconscious or non-"occurrent" belief, independent of a broader dispositionalist or functionalist program. Few of these accounts, however, appeal to phenomenal dispositions in their characterizations of belief, except secondarily and apologetically, with the promise that the phenomenal states themselves ultimately yield to functional or dispositional analysis.[12]

One reason for this difference is that previous dispositional accounts of belief have tended to serve a rather different purpose than the present account is meant to serve. Previous accounts have generally aimed to show how all talk about beliefs (in particular) or mental states (in general) can be transformed or "reduced" into talk about other, less objectionable things. Whether the motive is materialism or philosophical behaviorism, beliefs have seemed suspect – to be admitted only if they can be shown ultimately reducible to the tangible, the visible, or the material. Dispositional accounts have been offered as the key to such a reduction, if not practically speaking then at least in principle.

This project is no part of my own. I have no materialist or behaviorist agenda, and I aim at no such reductionist goal. The focus of the present account is on what tended to be only an afterthought to the thrill of exorcising the dualist's ghost-in-the-machine: a description of the conditions under which particular beliefs can properly be attributed to human beings. This account is meant to apply only to one case at a time, providing a framework with which to answer questions of this sort: When can we say of a person that she believes that P? Absent the reductionist agenda, one can answer such questions with no compunction about appealing to dispositions that involve other beliefs. Such appeals are circular only if one's project is to sketch a plan for the reduction of belief-talk as a whole.

Once a dispositional account of belief is unshackled from reductionist demands, the range of allowable dispositions broadens substantially. Not only can one appeal to other beliefs in one's dispositional characterizations, without any promise of an ultimate reduction to language free of belief, but one can unabashedly emphasize the importance of what I have been calling phenomenal dispositions – dispositions to undergo certain kinds of subjective, phenomenal experiences, such as the disposition to feel (not just exhibit) surprise and disappointment upon opening the fridge and finding one's beer gone. In labeling my account a "phenomenal" dispositional account, I mean to emphasize the role phenomenal dispositions play in belief.

Many of the anti-behaviorist objections to dispositional accounts of mental states are inapplicable as a consequence of the central role given to phenomenal dispositions in my account. The most compelling of these objections exploit the loose connection between mental states and behavior. Putnam (1963), for example, imagines a society of "super-spartans" who feel pain but do not exhibit the range of behaviors typically associated with pain (except perhaps avoidance, which is not specific to pain). Similarly, Strawson (1994) imagines a species of "weather watchers" who have beliefs and desires about the weather but are constitutionally incapable of acting in any way on the basis of those beliefs and desires. Chisholm (1957) emphasizes that we should not describe someone as disposed to act in a certain way, given a particular belief, unless we grant that that person has other particular beliefs and desires. For example, though Geronimo may believe that his aunt will be arriving at the railroad terminal in twenty-five minutes, it is only true to say he is disposed to go there to pick her up if he wants to pick her up and if his beliefs about how to get to the railroad terminal are not too deeply confused. Full conditions for the possession of any particular belief or desire can never be given in terms of behavioral dispositions alone; appeal to some other aspect of the subject's mental life will always be necessary.

The appeal to phenomenal dispositions gives the dispositionalist a clear and natural way around these objections. Putnam's super-spartans and Strawson's weather watchers, though they lack the manifest behaviors associated with believing, still have the phenomenal (and cognitive) dispositions characteristic of belief – if they did not, there really would be no reason to regard them as believing. Furthermore, they have clear, typical excusers from behavioral manifestation: contrary interests in the case of the super-spartans and incapacity in the case of the weather watchers. We can also grant Chisholm his point: There is no way to analyze away mental life in favor of behavioral dispositions or to replace all talk of belief with some other kind of talk. Since phenomenal dispositionalism does not aim to bring about these ends, it is no objection to phenomenal dispositionalism that it is impossible to do so.

I would like to conclude this section with some remarks about Ryle, the intellectual forefather of dispositionalism about mental states. Although he is typically viewed as a behaviorist for whom appeal to phenomenal dispositions would be strictly out of court (and he is, of course, committed to attacking the ghost-in-the-machine picture in a way that I am not), his case is more ambiguous than it first appears. Ryle certainly stresses the importance of behavioral dispositions and downplays the importance of phenomenal ones, sometimes even seeming to suggest that we could do without the latter entirely. Nevertheless, Ryle admits the relevance of such things as "silent colloquies" that others could not possibly overhear and tunes in one's head consisting of "the ghosts of notes similar in all but loudness to the heard notes of the real tune" (1949, pp. 184, 269). Noting such remarks, Hampshire (1950), in one of the earliest and most careful critiques of Ryle, regards him as having an "ambiguity of purpose" regarding the reduction of assertions about mental life entirely to statements about behavior. Despite his reputation, Ryle at times seems committed to the importance of internal, first-person phenomenology.

In light of this possibility, Ryle's short discussion of belief is interesting:

Certainly to believe that the ice is dangerously thin is to be unhesitant in telling oneself and others that it is thin, in acquiescing in other people's assertions to that effect, in objecting to statements to the contrary, in drawing consequences from the original proposition, and so forth. But it is also to be prone to skate warily, to shudder, to dwell in imagination on possible disasters and to warn other skaters. It is a propensity not only to make certain theoretical moves but also to make certain executive and imaginative moves as well as to have certain feelings (1949, pp. 134-135).

If we set aside for a moment the standard picture of Ryle as bent on reducing all talk about mental life to talk about behavioral dispositions, this passage begins to look rather like an appeal to a <u>mix</u> of behavioral, phenomenal, and cognitive dispositions. Perhaps a bit optimistically, then, I would like to claim Ryle as the first (albeit wavering) advocate of phenomenal dispositionalism about belief.

4. Mixed Sets of Dispositions.

The dispositional account of belief deals quite naturally with in-between cases of believing of the sort alluded to at the beginning of this paper, cases in which it seems not quite appropriate to describe the subject as either fully believing or not believing the proposition in question. In this section, I provide a few examples of such mixed cases and sketch some of the patterns into which they tend to fall.[13]

<u>4.1. Two Examples.</u>

Ellen, a twenty-year-old, studied Spanish for four years in high school. On the basis of her studies and her exposure to such Spanish words as 'mesa,' 'niña,' 'oreja,' and 'vaca,' she is willing, sincerely and cheerfully, to assent to the claim that all Spanish nouns ending in 'a' are feminine. Ellen has, however, occasionally come across certain

words ending in 'ista,' such as 'anarquista' and 'bolchevista,' that can be used either as masculine or as feminine (depending on the gender of the anarchist or Bolshevik), and she uses them correctly as masculine when the situation demands. She would not assent to the claim that all Spanish nouns ending in 'a' are feminine if an 'ista' word came to mind as a counterexample; nevertheless, in most circumstances she would not recall such counterexamples.

Does Ellen believe that all Spanish nouns ending in 'a' are feminine? Some of her dispositions accord with that belief; others do not. Whether it seems right to ascribe that belief to her varies contextually, depending on what dispositions interest us most. If we are considering which side she might take in a debate on the subject, it seems acceptable to say that she does believe that all Spanish nouns ending in 'a' are feminine. On the other hand, if we are interested in her skill as a speaker of Spanish and the likelihood of her making embarrassing gender errors in speech, it seems inappropriate to ascribe that belief to her. If we want to describe her doxastic state on the topic as carefully as possible, the best thing to do is to refuse to put the proposition 'all Spanish nouns ending in 'a' are feminine' either simply in or simply out of some imaginary "belief box" in her head, and instead to sketch the mix of her dispositions as I have just done.

Geraldine's teenage son Adam smokes marijuana. Usually Geraldine is unwilling to admit this to herself, and sometimes she adamantly denies it. Eating lunch with a friend, Geraldine can deplore her friend's parenting because of his daughter's drug use while denying in all sincerity that Adam has any similar problems. Yet she feels afraid and suspicious when Adam slouches home late at night with bloodshot eyes, and when she accuses him of smoking pot, she sees through his denials. In a certain kind of mood, she would tell her therapist that she thinks Adam smokes marijuana, but in another kind of mood she would genuinely recant such a confession. When Geraldine's husband voices concern about Adam's behavior, Geraldine sincerely comes to her son's defense. What does Geraldine believe on the subject? Someone insisting on a simple "Yes she

believes that he smokes marijuana" or "No, she doesn't" will be hard-pressed. Perhaps we could say that her beliefs on the subject change from situation to situation: When she is denying that her son smokes pot, she sincerely believes that he does not; when she is watching him creep in at 2:00 a.m., she sincerely believes that he does. But what does she believe <u>now</u>, while she's working intensely on a client's account and not giving the matter any thought? We may want to describe her as self-deceived, but even so, she is at most only <u>partially</u> self-deceived, since there are conditions under which she would unhesitantly acknowledge that her son uses marijuana.

The cases of Ellen and Geraldine are not meant to turn on any lack of knowledge on our part about their mental states, though lack of knowledge is a common source of hesitation in belief ascription. I am not putting forward an argument of the form: We cannot know what Ellen and Geraldine really believe; therefore, there is no fact about what they really believe. Rather, my suggestion is that we can know all there is to know about Ellen and Geraldine without being able to classify them as either genuinely believing or genuinely failing to believe the proposition in question. There is no simple answer to the question of what, "underneath it all," they really believe. Ellen and Geraldine are <u>between</u> believing and failing to believe the relevant propositions.

Although some cases of in-between believing become manageable simply upon recognition of <u>degrees</u> of belief, cases such as those described above do not yield to this approach. It is not that Ellen and Geraldine simply have a low degree of confidence (say .6 on a scale from 0 to 1) in the truth of the proposition in question. Rather, they are disposed to feel in some situations quite confident in asserting one thing, while at the same time they are disposed to feel in other situations quite confident in asserting the opposite. Their doxastic condition is far from the kind of simple uncertainty that one might feel, for example, about the outcome of an election or the toss of a die. The cases that are the focus of this paper are no more manageable by an analog view of belief, on

which belief is always smoothly describable by particular degrees of confidence, than by an all-or-nothing view of belief.

The dispositional account of belief recommends handling cases such as these by describing how the subject's dispositions conform to the stereotype for the belief in question and how they deviate from it. Subsequent questions may then be raised about the reasons for the match and mismatch of particular dispositions to the stereotype, opening avenues for both scientific research and everyday inquiry. Unlike at least some representational approaches to belief, the dispositional account of belief does <u>not</u> leave it as an open question whether, once their dispositional structure is fully characterized, Ellen or Geraldine really have the belief or not. There is no internal chalkboard on which the belief might be written in the language of thought, no warehouse in which it might be stored, despite its inconsistent "manifestations." Once all the relevant dispositions have been made clear, the case is closed. There are no further facts to report.

4.2. Normativity and Patterns of Deviation.

The usefulness of classifying people's mental states by appeal to stereotypical dispositional patterns depends on the tendency of people to adhere to these patterns. If cases such as Ellen's and Geraldine's were the norm, the dispositional stereotypes of belief would have little purpose. As a general rule, people who conform to some parts of the stereotype are apt to conform to other parts also. Furthermore, deviation from the stereotypes often falls into particular recognizable patterns, a few of which I will briefly sketch below.

The stereotypes capture more than merely statistical regularities, however. They capture something about how we think people <u>ought</u> to think, feel, and behave. Something about Ellen's and Geraldine's dispositional profiles strikes us as normatively lacking, as incoherent or confused. We feel that if Ellen and Geraldine correctly reasoned things through, they wouldn't deviate from the dispositional stereotypes in the way they

do. The conditional runs the other direction as well: Failures of reason will generally entrain failures to conform to the stereotypes.

This is not to say that conformity to all elements of the stereotypes is required by reason. For example, we can hardly convict someone of poor reasoning simply for not feeling disappointment upon suddenly learning that P, on which he had greatly counted, is false – strange though it may be in some cases and contrary to the stereotype. At the same time, however, something about such cases leaves us uneasy. Our folk psychology and everyday dealings with other people are so thoroughly dependent on the accuracy of these stereotypes that there is a kind of social accountability to the stereotypes that pervades even those aspects not shored up by the norms of reason. This, I think, is especially evident in the stereotypes associated with desires and personality traits, which are less thoroughly accountable to the strict demands of reason, and which consequently allow more room for social accountability to come undisguised into play. A person who is disposed greatly to enjoy ice cream on some occasions but to detest it on others, with no clear excusing conditions (such as detesting it only in times of grief or when she's already overfull with sweets), evokes this sort of discomfort. We want to know whether, "really," she likes ice cream or not, just as we want to know whether "really" Geraldine believes her son smokes marijuana. There will not always be in such cases a definite resolution of the kind we find satisfying. Still, we push for it: We want to fit people into our stereotypes, and there is social pressure on them actually to do so. In trivial cases, we tend not to be too deeply bothered, and the pressure is light. As the desires become more significant, or as deviation spreads across a larger variety of situations, the social reaction becomes more negative – and since such cases typically involve beliefs or moral values as well as desires, it can become unclear, even indeterminate, from which of a web of stereotypes it is most appropriate to describe the person as deviating.[14]

Nevertheless, certain patterns of deviation are pervasive enough that they don't at all strike us as strange, and in such cases we are much less likely to bring normative

pressures to bear. A person's motor behavior and expectations might accord with a certain belief stereotype, but not most of her inward and outward verbal dispositions, as might be the case, for example, with a driver who knows when and how far to turn the steering wheel to get into her driveway, but cannot express this knowledge in words. Similarly, people are often disposed to recognize and agree with assertions that P and to answer correctly a question like "P? Yes or no?" and yet not be able to come up with P as an answer to a more open-ended question or to act upon the truth of P when uncued. My dispositions regarding the last names of many of my acquaintances from college follow this pattern.

As a general rule, the more closely a mixed dispositional set matches a familiar pattern of deviation, the less puzzling it appears to us. At the other end of the spectrum are cases in which the subject's dispositions regarding P vary widely in no recognizable pattern at all. In the extreme, we would have to describe such cases as insanity.

A careful description of such in-between cases will tell exactly in what respects the subject deviates from the stereotype of the belief in question and in what respects the subject accords with that stereotype; it will look for a recognizable pattern in these deviations; and it will indicate which dispositions should count, in the present context, as the most important ones to the assessment. The description may or may not have a normative element of the sort described in this subsection.

4.3. A Short List of Patterns of Deviation.

It may be helpful to conclude this section by describing at least a few common patterns of deviation. The list below is by no means exhaustive. Some of the patterns of deviation on this list will suggest more irrationality on the part of the subject than others, in rough proportion to the extent to which the subject could, by simple reflection, bring herself in line with the stereotype.

Modularized believing: It is common for a subject's dispositional profile to match a belief stereotype in a narrow domain of practical expertise, but to deviate from the stereotype in most other domains and particularly with respect to the disposition to assent to the relevant propositions in speech. A driver's knowledge of how to steer into her driveway is an instance of this pattern. Cognitive psychologists sometimes describe such cases as cases of procedural (as opposed to declarative) knowledge.

Unconscious beliefs: Case studies in clinical psychology suggest that a subject may match a stereotype for believing that P in being disposed to claim that P under hypnosis or in free-association or in other of the techniques of psychotherapy; and the subject may exhibit hysterical or destructive symptoms that seem somehow consonant with a belief that P, though distorted; yet that subject may not be willing under normal circumstances to assent to P, even privately, because there is something unpleasant to the subject about the thought that P. This idea has been generalized into the popular notion of the unconscious, according to which a person may be disposed to act in a variety of ways in accordance with the stereotype for believing (or desiring) that P, yet because of the unacceptability of the thought that P, not be disposed to admit to himself that P is the case. Different people may assess differently the frequency of such cases, though it seems hard to deny that they at least sometimes occur.

Low confidence: People are not always perfectly confident in the truth or falsity of a proposition. Scholars interested in Bayesian decision theory have tended to characterize a subject's degree of confidence by associating a number from 0, indicating complete certainty in the falsity of the proposition in question, to 1, indicating complete certainty in its truth. If a person has modest confidence that P is true – a level of confidence that a Bayesian might characterize by saying that her degree of belief is .75 – it may be appropriate in some contexts to say, without qualification, that she believes that P, while in other contexts that simple ascription may be inappropriate or misleading. We may then qualify the belief ascription by saying that she suspects P to be the case or is inclined to

think that P, or by using Bayesian language. As a person's confidence declines, her match to the belief stereotype will also decline: She will be more likely to hesitate in asserting P, or may assert it only with qualification, or may refuse to assert it if the penalty for making a false statement is significant; she will feel less surprise if P turns out to be false; she will be less likely, unless she is interested in hypotheticals, to pursue the consequences of P very far in her reasoning; instead of offering that beer to the guest, Joe might say "let me check to see if we have any beer"; and so forth.[15]

Self-deception: Cases classified by folk psychology in the category of "self-deception" may be a subset of cases of unconscious believing. Geraldine's attitude toward her teenage son may fit, imperfectly, into this category of deviation. Philosophers have disputed at length about whether to describe the self-deceived subject as really believing the unpleasant proposition she would prefer to deny, or as really believing its opposite, or as somehow believing both. Few have endorsed, as I would, the view that the subject is not really accurately describable as believing either.[16]

Unreflective inconsistency: A subject may deviate from a stereotype simply because she fails to put two and two together. Ellen's case fits into this pattern. She matches the stereotype for believing that all Spanish nouns ending in 'a' are feminine in just those cases in which she is not reminded of a few exceptional nouns, and she deviates in cases in which those nouns become salient to her. We might suppose that with sufficient reflection, Ellen would come to match fairly exactly the stereotype for believing that not all Spanish nouns ending in 'a' are feminine. In cases of this sort, one would expect a match to the stereotype for believing a proposition in just those cases in which the reasons against believing that proposition are not salient.

Peripheral ignorance: A person may fail to match a stereotype due to ignorance of related topics. Examples of this pattern include the child who is uncertain about who the Pilgrims were and the case in which Joe believes that there is Budweiser in his fridge but does not believe that Budweiser is a type of beer. Hesperus-Phosphorus cases and

Kripke's (1979) case of Pierre, much discussed in the philosophy of language literature, may also be best describable as cases of in-between believing of this variety. Because Pierre does not know that 'London' and 'Londres' refer to the same city, he fails to conform completely to the stereotype for believing that London is ugly (e.g., he will deny the claim when it is made in French), although he does possess some dispositions central to the stereotype (e.g., the disposition to assert it sincerely in English). Everyday intuition seems to be fairly competent at determining what the dispositional effects of any particular type of peripheral ignorance might be.

Developing beliefs: This type of deviation is closely related to the previous two. Acquiring a network of knowledge in a particular domain and forging that knowledge into the kind of coherent structure necessary consistently to match the stereotype for various beliefs in that domain requires a certain amount of time. During this period of transition, the subject will have a mixed dispositional profile with respect to the relevant beliefs. According to Vygotsky (1978), for example, children typically acquire major new abilities and understandings by passing through a period during which they can exercise the knowledge or ability only with prompting or with proper structuring of the environment. As the child develops, less and less of this external "scaffolding" is necessary for the child to meet with success, and the child passes to fully developed competency.[17]

Partial forgetting: The processes of forgetting and unlearning, in some ways the opposite of belief development, also do not typically take place all at once. I am in the midst, for example, of forgetting the telephone number from my last apartment in Berkeley. Shortly after I left Berkeley, I could have rattled it off easily; later, it would have required more effort and sometimes would not have come; now I can recall the number only with a prompt of some sort; perhaps later I will be able to pick it out in a forced-choice test; in ten years, I may have no knowledge of it whatsoever. The more demanding the recall situation and the fewer the prompts provided, the less likely

someone in one of these intermediate stages of forgetting is to manifest the stereotypical behavior or phenomenology. This pattern of forgetting cannot typically be represented by a smooth decrease in degree of confidence. Rather, depending on the situation or the prompt, I may be very confident or completely uncertain about the truth of the same proposition. The psychological distinction between "recall" and "recognition" memory can to some extent capture the difference between the person who can easily rattle off a fact and the person who requires the fact to be presented before recognizing its truth, but even this is a rather crude way of splitting what is really a complicated gradation.

4.4. Conclusion of Section Four.

Talking about beliefs is useful because people with some of the dispositions in a stereotype will tend to have many of the other dispositions in that stereotype. Because of this, we can make generalizations and inductions on the basis of these stereotypes, and it is enormously convenient, even indispensable, to appeal to beliefs in describing our mental lives. Still, when there is a breakdown in the match between stereotype and the actual dispositional set of a subject, as will often happen in cases of the sort described above, simple belief talk may no longer be appropriate, and appeals to the stereotype may have to be replaced with more complicated appeals to specific dispositions or sets of dispositions or to recognizable patterns of deviation. On the account being offered in this paper, once the dispositional profile of the subject is made clear, it is a mistake to think that there is still some further question to be answered, namely, what does the subject really believe?

5. A Concern about Phenomenal Dispositionalism about Belief.

Functionalists such as Putnam (1966) and Lewis (1972, 1980), as well as some externalists about belief content such as Putnam (1975), Burge (1979), and Davidson

(1987) argue that the "content" of a belief is individuated not only in a forward-looking way, that is, by the phenomenology, behavior, and non-phenomenal mental states it is apt to produce, but also at least in part in a backward-looking way, by how it came about (or at least how states of its type are apt to come about). In other words, both groups of philosophers highlight the importance of looking back at the causes of beliefs in determining their content. In this section, I will address the concern that the dispositionalist account I propound will run against the arguments invoked in favor of the backward-looking elements in these positions.[18]

5.1. Externalism and phenomenal dispositionalism.

Externalists about belief hold that whether a subject believes that P, or whether the subject believes, instead, that Q, depends, at least sometimes, on facts about the world external to the subject herself. The dispositional account offered here is compatible with our intuitions in the kinds of cases typically invoked to support externalism. In fact, the present account comports more exactly with our intuitions in such cases than do standard externalist views.[19]

Consider Putnam's example of Twin Earth, a planet identical to Earth in every respect except that where Earth has water, Twin Earth has twater, indistinguishable from water by any of the tests available to the inhabitants of Earth or Twin Earth but in fact a different chemical compound than H_2O. Wayne from Earth and Dwayne from Twin Earth are molecule-for-molecule identical to each other (overlooking the water-twater difference). It seems intuitive to say that, despite the similarities between them, Wayne has beliefs about water, not twater, since that is what he interacts with on Earth, and Dwayne has beliefs about twater, not water (though both will, of course, use the word 'water' to describe what they see). If this is right, then it appears that the content of one's beliefs depends not only on what it is one's head, but also on one's environment and in particular on how one's beliefs were caused.

At first glance, it might seem that Wayne and Dwayne, being molecule-for-molecule identical to each other, could not possibly have different dispositions and thus must have the same beliefs on any dispositional account of belief. If this were so, then indeed the dispositional account of belief would run contrary to our intuitions in Twin-Earth-like cases. This would be unfortunate, perhaps, but not fatal: There is no guarantee that the most useful scientific or philosophical understandings of mind will accord with folk intuition in every respect.[20] As it turns out, however, dispositionalism about believing is compatible with such externalist intuitions, since dispositional properties may themselves be defined in part "externally," i.e., with reference to the organism's past or its environment. Only Wayne has the disposition to regard a present instance of water as an instance of the same kind of stuff Wayne drank as a child. Only Dwayne has the disposition to use the word 'water' intending to refer to the same kind of stuff to which people in his community on Twin Earth refer by using that word. Someone who believes that the meaning of a sentence is in part determined by factors external to the individual uttering that sentence has an additional pool of externally individuated dispositions to draw from in distinguishing Wayne from Dwayne. On that view, when Wayne utters the sentence 'water is clear and potable,' he is uttering a sentence that means water is clear and potable; when Dwayne makes exactly the same sounds, his sentence means that twater is clear and potable.

So there are at least some dispositions Wayne and Dwayne do not share. The question of whether we should describe them as having the same belief, then, depends on whether these differences are regarded as important enough in the context of ascription to warrant differential treatment of Wayne and Dwayne. If one chooses to focus on utterance meanings, and if these are individuated externally, or if one focuses on dispositions invoking one's past or one's community, one can fairly readily be drawn into regarding the men as having different beliefs. If one focuses instead on what it is like from the inside, on phenomenology and motoric behavior, and especially if one is not an

externalist about linguistic meaning, one may find oneself drawn in the internalist direction.[21] An internalist dispositionalist would hold that externally individuated dispositions are never relevant, for the purposes of philosophy or science, to the assessment of belief.

Although Putnam makes a good case for the intuitiveness of describing Wayne's and Dwayne's beliefs differently, in some contexts the intuitions are not so clear. For instance, let us suppose that Wayne and Dwayne are both environmental engineers working on a large water-treatment project. Miraculously, Wayne and Dwayne are teleported to each other's worlds. Wayne's coworkers may be concerned about Dwayne's ability to continue with the project. Doesn't it seem right to say that they shouldn't worry because Dwayne's opinions on the processes of water treatment are exactly the same as Wayne's?

Given that our intuitions on the Twin Earth case and other externalist cases described by philosophers such as Davidson (1987) and Burge (1979) are somewhat ambivalent and context dependent, as I think they are, the dispositional account of belief I have offered has an advantage over standard externalist accounts, since it provides room for such ambivalence and even allows us to predict contexts in which the intuitions may go one direction or the other. In the water-treatment case, the dispositions Wayne and Dwayne do share are the focus of concern, and so the dispositionalist account would predict an inclination to regard the two people as having the same belief. In other cases, where externally individuated dispositions are emphasized, the dispositional account will predict externalist intuitions. The dispositional account can thus accommodate and explain intuitions pulling in both directions, while standard externalist accounts must stand fast with an unchangeable answer: that what Wayne and Dwayne believe really is different. Standard externalists, therefore, are forced to try to explain away internalist intuitions the dispositionalist account handles quite naturally.

5.2. Functionalism and phenomenal dispositionalism.

Functionalists hold that what makes a state a belief with a certain content is its causal role in the system to which it belongs, or the causal role that states of its type typically play in systems of the type to which it belongs. Causal role has both forward-looking and backward-looking elements – something's causal role is determined both by what causes it or is apt to cause it, and by what it causes or is apt to cause. Pain is the favorite example. It is apt to be caused by, among other things, pinchings, pokings, heat, pressure, and bodily injury, and it is apt to cause, in turn, groaning, writhing, disrupted thought, and avoidance.

Although it is common for functionalists considering the individuation of mental states to argue for the importance of causal role generally, they typically run quickly over the question of whether in the functional analysis of belief one must include the backward-looking elements of causal role as opposed to only the forward-looking elements. Armstrong (1980) and Shoemaker (1981) both show, however, how such an argument would plausibly go. It begins with an attack on behaviorist dispositionalism like Chisholm's, discussed above: Because how one's beliefs dispose one to behave depends on one's desires, and how one's desires dispose one to behave depends on one's beliefs, it is impossible to reduce talk about mental states to any other kind of talk so long as one appeals only to behavioral dispositions. However, the typical functionalist, unlike Chisholm, shares the behaviorist's aim of analyzing mental predicates in terms of non-mental ones. Shoemaker says, for example,

> Let us say that a state (mental or otherwise) is functionally definable in the strong sense just in case it is expressible by a functional predicate that contains no mental predicates (or mental terminology) whatever.... It is functional states in this sense which functionalism takes mental states to be (1981, p. 95).

So long as one's task is to provide, for mental states, functional definitions in this strong sense, post-Rylean, anti-behaviorist arguments like Chisholm's show that mere appeal to forward-looking dispositions will not do. Armstrong and Shoemaker appeal, therefore, not only to dispositions to behave but also to the typical physical causes of mental states and to the causal relations between mental states, on the hope that the whole bundle of mental states, taken together, can in principle be characterized wholly in terms of physically (or at least non-mentally) described inputs and outputs. Since it is not part of the project of phenomenal dispositionalism to characterize mental predicates by means of non-mental predicates, however, the functionalist's reasons for wanting to appeal to backward-looking causal relations in individuating mental states do not apply.

It may be that there is some warrant for a revised functionalism that characterizes and individuates mental states both dispositionally <u>and</u> in terms of how they are apt to come about, but at the same time does not require that mental predicates be in principle characterizable by non-mental ones – a functionalist account, in other words, that does not treat phenomenal (and other mental) states as simply as falling out of the functional relations but rather treats them as among the fundamental relata. I have no serious objections to such a view, although in the case of belief in particular I am inclined to make the stronger claim that once one takes phenomenal dispositions as fundamental, an adequate characterization of what it is for a subject to believe something does not require appeal beyond the subject's forward-looking dispositions. To argue otherwise would require quite a different set of objections than can readily be drawn from the functionalist literature.

6. Conclusion

I hope that the preceding account of belief is a useful one. It does not give a simple, straightforward criterion for assessing the presence or absence of a belief, but no account

that gives such a criterion accords at all well with our intuitions about the conditions under which it is appropriate to attribute beliefs; and for the present at least it seems desirable to have some sort of account that matches reasonably well with those intuitions. The account is no more complex, and quite possibly a fair bit simpler, than functionalist approaches, which must appeal to the same sorts of <u>ceteris paribus</u> generalizations as the present account does, and internal representation approaches, which, if interpreted realistically, posit covert entities related in a complex way with anything directly observable (at least to a third party – and plausibly, I think, even in one's own case). The account may also avoid some of the difficulties that beset other accounts, though no doubt it brings difficulties of its own that I do not have the foresight to guess. I have argued that the account is especially useful, while other approaches are at best vague, in handling "in-between" cases of believing of the sort that should interest anyone concerned with the messy details of development, forgetting, irrationality, ignorance, context-dependence, or other such issues.

The metaphysically inclined may wish to ask whether, useful or not, the present account accurately describes what belief <u>really</u> is. I must admit that I fail to feel the impulse that drives questions such as this – and corresponding questions in other areas of philosophy, such as, What is a person, a cause, or free will, <u>really?</u> There are useful and less useful ways to think about such topics, ways that correspond better with divisions and tendencies of the sort reported by empirical sciences, ways that promote or hinder a particular vision of human flourishing and the development of moral community. A philosophical account or concept may prove useful in one context or relative to one set of goals and a hindrance in another context or with other goals. As far as I can see, there is no more sense in the insistence that one account or another is the real, honest, metaphysical truth about things than there is in insisting that an ace is really, honestly, metaphysically the highest card, independently of the game that is being played.

Representational approaches to belief have played an important role in cognitive science, and this account is not meant to displace them, but to supplement or complement them. In some contexts, I believe, the present account will prove more useful; in others, a representational approach will work as well or better. The likelihood that the accounts will in some cases give contradictory answers to questions about belief implies that one cannot think of both accounts as literally true and universally applicable. But this need be no hindrance to endorsing them conditionally or as idealizations – and we endorse (or should endorse) most scientific theories no differently (see Cartwright 1983, 1999 and Dupré 1993).

Someday, an advanced neuroscience may replace our talk of belief (as envisioned in Churchland 1981) or discover, corresponding to each of our beliefs, specific neural bases of the sort envisioned in the most optimistic type-type reductionism. Or, less dramatically, the advance of knowledge may force us to revise substantially the concept of belief, beyond the modest kind of revisions suggested here and elsewhere in the philosophical literature.[22] If the type of dispositional approach to belief offered here is thereby rendered obsolete, that in no way vitiates its usefulness now.[23]

[1] A few of the most obvious advocates of this view include Fodor (e.g., 1990), Dretske (1988), Millikan (1984), Cummins (1996), and Searle (1983).

[2] Although he does not cast his concerns as concerns for the metaphysics of representations, some of these complications appear as concerns about clutter avoidance and conditionalization in Harman (1986). For example, the well-known difficulty of how to account on a representational theory for the apparently infinite number of beliefs each of us has (e.g., I believe that the number of planets is less than 10, also that it is less than 11, etc.) becomes even more intractable if propositions about which one has no settled opinion are beliefs of low or intermediate degree of confidence. It may then follow that for every possible proposition, one must have a representation of some sort in mind. Also, philosophers such as those described in the previous note have struggled to describe naturalistically how a belief gets its content and the role it plays in action. This task is apt to become much more complicated by the introduction of degrees of belief: If (simplifying a bit) the belief that a cow is there is just the belief apt to be caused, in normal circumstances, by a cow's being there, what are we to say about the belief, with .4 degree of confidence, that a cow is there?

[3] Schwitzgebel (2001) develops this point in a little more detail and explores some novel cases of in-between believing. However, most of the ideas appearing in that short paper are developed in much greater depth in the present work.

[4] Putnam (1975) offers a similar view of stereotypes. I differ from Putnam in associating stereotypes with things, classes, and properties instead of words and in seeing stereotypes as clusters of properties rather than sets of ideas.

[5] For a review of the literature on dispositions, see Prior (1985). A spread of views can also be found in Tuomela (1978) and Armstrong, Martin, and Place (1996).

[6] Since I will be characterizing belief in terms of stereotypical dispositional properties, variations of this sort, especially in extreme cases, may lead to a degree of relativism about the accuracy of belief ascriptions: It may turn out that for you, with your stereotypes, it is accurate to describe me as having a particular belief, while for someone else, with her stereotypes, it is not accurate to describe me that way. Although I suspect that differences in stereotype will tend not to be large enough to produce substantial differences in the appropriateness of belief ascriptions, relativism of this sort is not in any case as odious as it may at first appear. On the view espoused in this paper, dispositional facts are fundamental in matters of belief, and the language of belief is employed as a convenient way of grouping together dispositional properties that tend to co-occur. Since the accuracy of attributions of dispositional properties to a particular person does not vary from ascriber to ascriber (caveat: see note 9), the relativism about belief that may be engendered by my account is ontologically superficial – it is a consequence only of variations in people's shorthand ascriptions and does not reflect real differences in the subjects themselves.

[7] It may be that some of the some of the stereotypical dispositions that I have presented as behavioral are really tacitly hybrid dispositions of this sort. How exactly to work this out would depend on one's view of tacit conditions of manifestation (see section 2 and note 9) and whether, for example, the disposition deliberately to walk to the fridge is equivalent to the disposition to walk to the fridge, given certain tacit conditions of manifestation.

[8] Cases for the usefulness and non-vacuity of ceteris paribus laws and generalizations are made by Cartwright (1983) and Pietroski and Rey (1995). The river example is from Fodor (1987), who uses it for similar purposes.

[9] One metaphysical question that arises in this regard is whether two dispositions are identical if they differ only in that one has as a tacit condition what the other has as an

explicit condition. Treating two such dispositions as identical might be motivated by the thought that the difference between the tacit and explicit conditions is only in how precise we are in characterizing the disposition in question. If one accepts both this view and the view that dispositions have an indefinitely large number of tacit conditions, then one might think of dispositions as not so much ceteris paribus defeasible as too complicated to specify completely. If, alternatively, one chooses to individuate dispositions entirely on the basis of their explicit conditions, exactly what dispositions the stereotypes contain will be contingent upon how widespread certain conditions are, since if a belief, desire, practice, or environmental condition is widespread, it may appear in the tacit rather than the explicit conditions of manifestation of the dispositions in the belief stereotypes to which it is relevant. A third position joins the view that dispositions need not always manifest when their conditions of manifestation are met with the view that the difference between tacit and explicit conditions is irrelevant to the individuation of dispositions. This view can also lead to contingency in the contents of belief stereotypes, if it is contingent whether a particular background condition necessary for the manifestation of a stereotypical disposition is included among the conditions of manifestation, whether tacit or explicit. Such ontological quibbles, however, need have no impact on belief assessment, even if the contents of the stereotypes are nominally different, since their main practical difference is just in the characterization certain permitted failures of dispositional manifestation – as, for example, excused deviations rather than consequences of unmet implicit conditions.

[10] When considering a disposition for the purposes of assessment of adherence to a belief stereotype, the evaluation of excuses may be different than when considering, for other purposes, a disposition with the same manifestation and explicit conditions. Physical incapacity to utter the required phonemes will not necessarily be a ceteris paribus excuser from the disposition to utter "there is a beer in my fridge" if we are

interested in that disposition in grading Joe's tongue function. A difference of this sort in allowable excuses will show up in the tacit conditions of manifestation of a disposition, since allowable excuses is just what the tacit conditions of manifestation indicate.

[11] The classic treatment of mental states as dispositionally specifiable is in Armstrong (1968). A few of the more prominent functionalist treatments include Lewis (1972, 1980), Putnam (1966), and Fodor (1968). Braithwaite (1932-1933) anticipates some important elements of Ryle's approach to belief.

[12] Searle is an exception to this tendency, but his dispositional account is exclusively of unconscious beliefs, which he sees as derivative of a more basic sort of occurrent, conscious belief.

[13] Stich (1983) is a good source of further examples, although Stich does not endorse a dispositional account of belief. Dennett (1987), especially pp. 103-116, is also interesting on this issue, and his view is in many respects similar to my own.

[14] The ideas in this paragraph were developed in conjunction with Tori McGeer. Related ideas are developed in McGeer (1996) and Brandom (1994).

[15] One might be tempted to the view that any proposition held with greater than .5 confidence is genuinely believed (albeit weakly). However, it is no more necessary to endorse the view that a subject who believes P with .55 confidence (and thus believes not-P with .45 confidence) genuinely believes that P than to endorse the view that some color that is 25% blue, 20% red, and 55% yellow (on some scale) is genuinely yellow.

[16] A good review of the literature on self-deception can be found in Mele (1987b). Mele (1987a), Barnes (1997), and Lazar (1999) espouse the view that the self-deceived person really believes what she claims to believe. The view that the self-deceived person really believes the contrary of what she avows is defended by, among others, Bach (1981) and Audi (1982, 1985). The view that the self-deceived subject really has both beliefs

has been defended by Rorty (1972, 1988), Pears (1984), Da Costa (1990), and Davidson (1998).

[17] Such cases are explored in more detail in Schwitzgebel (1999).

[18] A related concern that has frequently been raised in conversation against my account is that it doesn't allow for beliefs to cause behavior. While a full treatment of this concern would require a lengthy expedition into the metaphysics of dispositions, a little can be said here. First, I am not committed to Ryle's view of dispositions as bare conditionals. Perhaps it is plausible to think that underlying each disposition is a "categorical basis" causally responsible for the dispositional manifestation when the conditions of manifestation are satisfied. Armstrong (1968) identifies each disposition with its categorical basis (but this view may be weakened later in Armstrong, Martin, and Place 1996), but more complicated views may also be plausible (e.g., in Prior 1985). If a disposition just is its categorical basis, then dispositions, by definition, cause their manifestations. Even if the relation between disposition and categorical basis is complicated, it is not unreasonable to suppose that in every case the dispositional profile of a believing subject will have a categorical basis underlying it that is causally responsible for the dispositional manifestations when they occur. This categorical basis may be identified with the belief in question. It then follows that the belief causes the dispositional manifestations. I am willing to allow the identification of believing with being in a certain categorical state as long as that state co-occurs, in all nomologically possible worlds, with the appropriate dispositional profile. If deep metaphysical reasons exist, maybe there are some such reasons for putting such a categorical spin on things; but as I have stipulated the case, it is difficult to imagine any nomologically possible evidence for a belief dispositionally characterized that is not equally evidence for that belief characterized categorically. It thus seems unlikely that the proponent of the categorical move would disagree with with a dispositional purist about the appropriateness of any

actual belief ascription. When Armstrong and the functionalists convinced much of the philosophical community to identify mental states not with bare behavioral dispositions but with the categorical states typically underlying them, part of the appeal of that move was that it allowed people to ascribe mental states even when the subject was not actually disposed to manifest typical behavior. That is an improvement over behaviorism, but I have argued that an analogous disconnection is not an improvement over my account: Once the dispositions are fully characterized, the question of what the subject believes is closed.

[19] I will not discuss what Hurley (1998) calls "vehicle externalism" (in contrast to "content externalism"). I see no reason to think it incompatible with a dispositional approach to belief.

[20] Fodor's (1981) position of "methodological solipsism" (an idea he borrows and expands from Putnam 1975), for example, is the view that although something like the folk concepts of belief, desire, etc., may be appropriate for psychological theorizing about the mind, these concepts must be purged of their externalist consequences. Fodor (1994) revises or abandons this view. For more on my view on the possibility of revising folk conceptions for scientific purposes, see Gopnik and Schwitzgebel (1998).

[21] Dretske (1995) argues, however, that even phenomenology varies externalistically.

[22] Alison Gopnik and I argue for the likelihood of such changes in Gopnik and Schwitzgebel (1998).

[23] Dozens of people have helped me refine the ideas that are the subject of this paper. I would like especially to thank Max Deutsch, Josh Dever, Martin Jones, Alison Gopnik, Ariela Lazar, Lisa Lloyd, William Lycan, Kim Kempton, John Madsen, Maria Merritt, Chris Olsen, Pauline Price, John Searle, and Leo van Munching. I owe Tori

McGeer and John Heil a particular debt, as it was in conversation with them that I first

became convinced that a dispositional account of belief could work.

REFERENCES

Armstrong, D. M. (1980) <u>The Nature of Mind</u>, St. Lucia, Queensland: Queensland.

Armstrong, D. M. (1968) <u>A Materialist Theory of Mind</u>, London: Routledge and Kegan Paul.

Armstrong, D. M., Martin, C. B., and Place, U. T. (1996) <u>Dispositions</u>, London: Routledge.

Audi, Robert. (1982) "Self-Deception, Action, and Will," <u>Erkenntnis</u> 18, pp. 133-158.

Audi, Robert. (1985) "Self-Deception and Rationality," in M. W. Martin (ed.), <u>Self-Deception and Self-Understanding</u>, Lawrence, KA: University Press of Kansas.

Audi, Robert. (1994) "Dispositional beliefs and dispositions to believe," <u>Nous</u> 28, pp. 419-434.

Bach, Kent. (1981) "An Analysis of Self-Deception," <u>Philosophy and Phenomenological Research</u> 46, pp. 351-370.

Barnes, Annette. (1997) <u>Seeing Through Self-Deception</u>, Cambridge: Cambridge.

Braithwaite, R. B. (1932-1933) "The Nature of Believing," <u>Proceedings of the Aristotelian Society</u> 33, pp. 129-146.

Brandom, Robert. (1994) <u>Making It Explicit</u>, Cambridge, MA: Harvard.

Carnap, Rudolf. (1950) <u>Logical Foundations of Probability</u>, Chicago: University of Chicago.

Cartwright, Nancy. (1983) <u>How the Laws of Physics Lie</u>, New York: Oxford.

Cartwright, Nancy. (1999) <u>The Dappled World</u>, Cambridge: Cambridge.

Chisholm, Roderick M. (1957) <u>Perceiving: A Philosophical Study</u>, Ithaca, NY: Cornell.

Churchland, Paul M. (1981) "Eliminative materialism and the propositional attitudes," <u>Journal of Philosophy</u> 78, pp. 67-90.

Cummins, Robert. (1996) <u>Representations, Targets, and Attitudes</u>, Cambridge, MA: MIT.

Da Costa, Newton C. A. (1990) "Belief, Contradiction, and the Logic of Self-Deception," American Philosophical Quarterly 27, pp. 179-197.

Davidson, Donald. (1987) "Knowing one's own mind," Proceedings and Addresses of the American Philosophical Association 60, pp. 441-458.

Davidson, Donald. (1998) "Who Is Fooled?" in J.-P. Dupuy (ed.), Self-Deception and Paradoxes of Rationality, Stanford, CA: CSLI.

Dennett, Daniel C. (1987) The Intentional Stance, Cambridge, MA.: MIT.

Dretske, Fred. (1988) Explaining Behavior, Cambridge, MA: MIT.

Dretske, Fred. (1995) Naturalizing the Mind, Cambridge, MA: MIT.

Dupré, John (1993) The Disorder of Things, Cambridge, MA: Harvard.

Fodor, Jerry A. (1968) Psychological explanation, New York: Random House.

Fodor, Jerry A. (1981) Representation, Cambridge, MA: MIT.

Fodor, Jerry A. (1987) Psychosemantics, Cambridge, MA: MIT.

Fodor, Jerry A. (1990) A Theory of Content, Cambridge, MA: MIT.

Fodor, Jerry A. (1994) The Elm and the Expert, Cambridge, MA: MIT.

Gopnik, Alison, and Schwitzgebel, Eric. (1998) "Whose Concepts Are They, Anyway? The Role of Philosophical Intuition in Empirical Psychology," in M. R. DePaul and W. Ramsey (eds.), Rethinking Intuition, Lanham: Rowman & Littlefield.

Hampshire, Stuart. (1950) "The Concept of Mind," Mind 59, pp. 237-255.

Harman, Gilbert. (1986) Change in View, Cambridge, MA: MIT.

Hurley, Susan. (1998) Consciousness in Action, Cambridge, MA: Harvard.

Kripke, Saul A. (1979) "A Puzzle about Belief," in A. Margalit (ed.), Meaning and Use, Dordrecht: Reidel.

Lazar, Ariela. (1999) "Deceiving Oneself or Self-Deceived? On the Formation of Beliefs 'Under the Influence,'" Mind 108, pp. 265-290.

Lewis, David. (1972) "Psychophysical and Theoretical Identifications," Australasian Journal of Philosophy 50, pp. 249-258.

Lewis, David. (1980) "Mad Pain and Martian Pain," in N. Block (ed.), Readings in the philosophy of psychology, vol. 1, Cambridge, MA: Harvard.

Marcus, Ruth B. (1990) "Some Revisionary Proposals about Belief and Believing," Philosophy and Phenomenological Research 50, pp. 133-153.

Martin, C. B., and Heil, John. (1998) "Rules and Powers," Philosophical Perspectives 12, pp. 283-312.

McGeer, Victoria. (1996) "Is Self-Knowledge an Empirical Problem? Renegotiating the Space of Philosophical Explanation," Journal of Philosophy 93, 483-515.

Mele, Alfred R. (1987a) Irrationality, New York: Oxford.

Mele, Alfred R. (1987b) "Recent Work on Self-Deception," American Philosophical Quarterly 24, pp. 1-17.

Millikan, Ruth. (1984) Language, Thought, and Other Biological Categories, Cambridge, MA: MIT.

Pears, David. (1984) Motivated Irrationality, Oxford: Clarendon Press.

Pietroski, Paul, and Rey, Georges. (1995) "When Other Things Aren't Equal: Saving Ceteris Paribus Laws from Vacuity," British Journal for Philosophy of Science 46, 81-110.

Prior, Elizabeth. (1985) Dispositions, New Jersey: Aberdeen.

Putnam, Hilary. (1963) "Brains and Behavior," in R. Butler (ed.), Analytical Philosophy, Oxford: Basil Blackwell.

Putnam, Hilary. (1966) "The mental life of some machines," in H.-N. Castañeda (ed.), Intentionality, Minds, and Perception, Detroit: Wayne State.

Putnam, Hilary. (1975) "The Meaning of 'Meaning,'" in K. Gunderson (ed.), Minnesota Studies in the Philosophy of Science, 7: Language, Mind, and Knowledge.

Rorty, Amelie O. (1972) "Belief and Self-Deception," Inquiry 15, pp. 387-410.

Rorty, Amelie O. (1988) "The Deceptive Self: Liars, Layers, and Lairs," in B. P.

> McLaughlin and A.O. Rorty (eds.), <u>Perspectives on Self-Deception</u>, Berkeley, CA:
> University of California.

Ryle, Gilbert. (1949) <u>The Concept of Mind</u>, New York: Barnes and Noble.

Schwitzgebel, Eric. (1999) "Gradual Belief Change in Children," <u>Human Development</u>

> 42, pp. 283-296.

Schwitzgebel, Eric. (2001) "In-Between Believing," <u>Philosophical Quarterly</u> 51, pp. 76-

> 82.

Searle, John. (1983) <u>Intentionality</u>, Cambridge: Cambridge.

Shoemaker, Sydney. (1981) "Some Varieties of Functionalism," <u>Philosophical Topics</u> 12,

> pp. 93-119.

Stich, Stephen. (1983) <u>From Folk Psychology to Cognitive Science</u>, Cambridge, MA:

> MIT.

Strawson, Galen. (1994) <u>Mental Reality</u>, Cambridge, MA: MIT.

Tuomela, Raimo (ed.). (1978) <u>Dispositions</u>, London: Routledge.

Vygotsky, Lev S. (1978) <u>Mind in Society</u>, M. Cole, V. John-Steiner, S. Scribner, and E.

> Souberman (eds.), Cambridge, MA: Harvard.