

**A Dispositional Approach to Attitudes:
Thinking Outside of the Belief Box**

Eric Schwitzgebel

Department of Philosophy

University of California at Riverside

Riverside, CA 92521-0201

USA

eschwitz at domain- ucr.edu

November 1, 2012

A Dispositional Approach to Attitudes: Thinking Outside of the Belief Box

1. Introduction

I offer here an account of what it is to have an attitude. I intend this account to be entirely general – to include all the “propositional attitudes” (believing, desiring, intending, fearing, hoping...), the “reactive attitudes” (resenting, appreciating, forgiving, being angry with...), and other types of attitudes that appear to be directed toward people, things, or events (loving Tim, trusting Elena, hating jazz, having a “bad attitude” about school, valuing kindness over intelligence, approving of the President’s foreign policy decisions...). I will argue that to have an attitude is, primarily, (1) to have a dispositional profile that matches, to an appropriate degree and in appropriate respects, a stereotype for that attitude, typically grounded in folk psychology, and secondarily, (2) in some cases also to meet further stereotypical attitude-specific conditions. To have an attitude, on the account I will recommend here, is mainly a matter of being apt to interact with the world in patterns that ordinary people would regard as characteristic of having that attitude.

One alternative view – perhaps the dominant alternative view – treats having an attitude as a matter of possessing some particular internally stored representational content, a content perhaps poised to play some specific set of cognitive roles depending on the attitude type (e.g., Millikan 1984; Fodor 1987; Dretske 1988; Cummins 1996; Carruthers 2006). On such a view, it is sometimes said, to believe that P (for example, to believe that snow is white) is to have a representation with the content “P” (“snow is white”) stored in a metaphorical “Belief Box”; to

desire that Q is to have a representation with the content “Q” stored in a metaphorical “Desire Box”; etc. This type of metaphor is, I believe, misleading – partly for reasons that will emerge below. (See this note¹ for some further considerations.)

The term “attitude” once meant *posture*, especially in statuary and painting, deriving from the Latin, French, and Italian terms for aptitude or fitness or disposition. Only in the 19th century did the word acquire its psychological use.² I recommend that we retain a sense of those etymological origins. An attitude is a temporary or habitual posture of the mind. The tools of everyday folk psychology provide us with ways of labeling typical postures, and a person’s own posture may match such typical postures to a greater or lesser extent.

2. *A Motto and Some Examples.*

Motto: To have an attitude is, at root, to live a certain way.

To regard one’s colleagues as a talented group is, at root, to live a certain way. It is, in part, to be disposed to affirm – when appropriate circumstances arise – that one’s colleagues are a talented group. But more than that, it is to regard many or most of one’s colleagues, considered individually, as talented. It is to feel a certain pride to be among them (perhaps a slightly embarrassed pride, if one sees oneself as not quite on par). It is to be unsurprised when they collect honors. It is to think of them as natural choices to contribute to compilations in their areas of interest. It is to tend to expect them to make worthwhile colloquium comments, to seek

¹ See critiques of the warehouse model of memory in Bartlett 1932; Neisser 1967; Roediger 1980; and Sutton 2008; and my critique of such models in developmental psychology in Schwitzgebel 1999 and McGeer and Schwitzgebel 2006.

² See the *Oxford English Dictionary* entry for “attitude” (<http://www.oed.com>). The phrase “posture of the mind” I owe to Annette Baier, who launches her 1985 book, *Postures of the Mind* by noting and endorsing its anti-atomistic character. Baier in turn credits Locke 1690/1975.

them out for insight, and to feel something important has been lost should they leave the department. It is to be hesitant to draw the conclusion that their oral and written remarks are foolish. And so forth – at least *ceteris paribus* (all else being equal or normal or right; see Section 4 below). It is to have, in general though probably only imperfectly, a certain profile of outward behavior and inner experience – a profile recognizable in broad outline to non-academics and in more specific detail to members of one's own discipline. To have this attitude is to embody a certain broad-ranging actual and counterfactual pattern of activity and reactivity.

To value one's marriage is likewise, at root, to live a certain way. It is in part to be disposed to say that one values one's marriage. But more than that, it is to ungrudgingly invest time in working through tensions. It is not seriously to consider activities, such as extramarital affairs, that might threaten the marriage. It is to be willing to compromise one's career. It is to be disposed to react with alarm to an apparent shift in one's spouse's attitude. To value one's marriage is to have a certain general psychological posture toward one's marriage, reflected broadly across a wide range of actual and potential thoughts, emotions, and actions. When your spouse rebukes you for insufficiently valuing the marriage, it is for your failure to fit this pattern.

To love baseball, too, is to live a certain way. It is to enjoy watching and participating in baseball games, to leave room for baseball in one's plans, to talk baseball with other aficionados, to relish the onset of the season, to care intensely about the outcome of certain games, and so forth – or at least to be disposed in most of these directions, *ceteris paribus*. Similarly for disliking school, fearing nuclear war, intending to move to Hawaii, being satisfied with one's life, having racist or sexist or patriotic attitudes, believing in God, wanting fame, forgiving an insult, hoping to play Road Kill Rally with Dan, and disapproving of cigarettes. As the attitudes become narrower – believing that one's daughter is wearing red sandals, wanting a plum rather

than a peach – it perhaps sounds grandiose to say one “lives that way”. But that’s because the broad pattern of action and reaction in such cases drifts mostly into the realm of the counterfactual; counterfactually there still is a vast pattern of activity around preferring the plum.³

I think we have – not only as theoreticians but also as ordinary people engaged in personal interaction – some degree of choice about how to talk about and theorize the attitudes. Perhaps you will find something attractive in the approach gestured at above.

3. *Deep vs. Superficial Accounts of Psychological States.*

Let’s say that relative to a class of *surface phenomena*, an account of a property is *deep* if it identifies possession of the property with some feature other than patterns in those same surface phenomena – some feature that explains or causes or underwrites those surface patterns. In contrast, let’s say that an account is *superficial* if it identifies possession of the property simply with patterns in the surface phenomena.

Consider the property of *being a cat*. Here’s one potential class of surface phenomena relative to which an account of cat-hood might be deep: the class of currently macroscopically observable properties of the species, such as having four legs, having fur, having a heart and kidneys, and tending to chase and eat certain types of smaller animals. A superficial account of what it is to be a cat might identify being a cat with possession of some cluster of such surface phenomena. In contrast, a deep account will appeal beyond this cluster, to properties that

³ Might an attitude never actually manifest itself in any way, but only counterfactually? I don’t want to exclude that possibility for marginal cases. However, our most important attitudes, the ones I want to treat as paradigmatic, are the attitudes that actually show in our face and limbs and reasoning. Against insufficiently lived attitudes, see also Section 8 on intellectualism and Section 11 on wraiths.

presumably explain the existence of the surface properties – properties such as having a certain DNA structure or having a certain evolutionary history. When the surface phenomena and the deep properties diverge, surface accounts and deep accounts will disagree about classification. For example, if an entity molecule-for-molecule identical to a cat were spontaneously to congeal, by freak chance, from relatively disorganized matter, such an entity would still qualify as a cat on a superficial account but not on a deep account that requires that the entity have a particular evolutionary history.⁴ It might be useful to possess both superficial *and* deep accounts of cat-hood, even different kinds of superficial and different kinds of deep accounts that differently draw the line between the surface and the depths. Deep accounts are not always better or more scientific.⁵

Accounts of psychological properties can likewise be deep or superficial relative to a class of surface phenomena. Any account of a psychological property that identifies possession of that property with being in a particular brain state, for example, will be deep relative to any class of surface phenomena that does not include brain states. Any account of a psychological property that identifies possession of that property with having a particular folk-psychologically non-obvious functional architecture will be deep relative to any class of surface phenomena that does not include folk-psychologically non-obvious functional architecture. In both of these respects, my approach to the attitudes is superficial rather than deep.⁶

4. Basics of the Approach.

⁴ A “Swampcat” case: Davidson 1987; Dretske 1995; Millikan 1996; Neander 1996.

⁵ Cf. Dupré 1993, 2002.

⁶ Compare symptom-based vs. etiological or physiological approaches to physical and mental disease (Murphy 2006).

I recommend that we embrace the following two-clause claim:

(1.) To have an attitude is, primarily, to have a dispositional profile that matches, to an appropriate degree and in appropriate respects, a stereotype for that attitude, typically grounded in folk psychology.

(2.) Some attitudes also require meeting further stereotypical attitude-specific conditions.

Claim 1 is the heart of the approach. I offer Claim 2 somewhat in the spirit of concession, as I will explain below.

4.1. Dispositions. Claim 1 employs the concept of a dispositional profile. A dispositional profile is a suite of dispositional properties, or more briefly dispositions. I won't enter into the large literature on dispositions, since I think my account can fit with a variety of approaches.⁷ However, I like to think of dispositions as modal generalizations: They concern not primarily what *actually* occurs, or what *would occur in some particular instance*, but rather what *would occur across a range of relevant, normal conditions*. For example, to be disposed to interpret one's colleagues' remarks charitably is to be such that, counterfactually, across a range of normal circumstances, one would interpret one's colleagues' remarks charitably.

As Ryle (1949) has emphasized, dispositional properties come in different flavors. For example, there are tendencies, capacities, liabilities, and pronenesses, each differing in modal profile. Someone might have the *capacity* to do something without the *tendency* to do it, for example, the capacity to invest time in smoothing marital tensions without the tendency to do so.

⁷ E.g., Ryle 1949; Prior 1985; Armstrong, Martin, and Place 1996; Mumford 1998; Fara 2005. It would be convenient for my view if dispositions caused their manifestations, but I am not committed to that. If the vase's fragility causes it to break when dropped, so also does the belief that P cause the assertion that P. If not, we need an error theory in both cases. I would also consider the metaphysical jujitsu move of simply identifying the attitude, token-token, with whatever happens to be the categorical basis of the relevant bits of the individual's current dispositional structure, if that delivers the desired facts about causation.

Conversely, someone might have a tendency without a capacity: Someone might tend to make arithmetic mistakes when listening to music but it seems odd to say that's a capacity she has. Liabilities also differ from pronenesses: A vase might be liable to break without being likely enough to break for us to want to say it is prone to break. Etc. I emphasize this point because different attitudes appear to require different types of dispositional property. For example, love might require actually manifested tendencies and not merely unmanifested pronenesses.⁸ Knowledge might be better conceptualized as involving capacities than tendencies.⁹

Dispositions generally manifest only *ceteris paribus* – all else being equal or right or normal, or absent a countervailing force. I might be disposed to interpret my colleagues' remarks charitably and yet not actually interpret their remarks charitably – if I have fallen asleep, for example, or if I'm in a foul mood. Perhaps in a majority of actual cases and in a substantial portion of nearby counterfactual space I don't interpret their comments charitably. As long as such non-manifestations are, shall we say, *excused* by the fact that not all else is right or normal, it can still be true that I have the disposition. The same *ceteris paribus* defeasibility is true of most scientific generalizations, as emphasized by Nancy Cartwright (1983). Modal generalizations – generalizations that are not simply summations across actual instances – typically operate across an implicitly assumed background in which things are normal and competing forces negligible.

4.2. *Stereotypes*. A stereotype for a property X is a cluster of other properties that would be regarded as characteristic of something that possesses property X. Some dispositions might

⁸ Perhaps in this way “dispositional” (e.g., Naar forthcoming) and “historicist” or “emotion complex” (e.g., Helm 2005/2009) accounts of love can be partly reconciled.

⁹ See Ryle 1949; Margolis 1973; Myers-Schulz and Schwitzgebel forthcoming.

be more central to the cluster and others more peripheral; and the contributing dispositional elements might not be entirely discrete.

A property might be *regarded* as characteristic even if it's not actually characteristic. It might be stereotypical of fashion models, and yet false on the whole, that they lack verbal wit. And to say that a property *would* be regarded as characteristic is not to say that it is *actually* regarded as characteristic. Maybe no one has ever explicitly considered what properties would be characteristic of a nerdish love of squids. But we can readily enough attribute characteristic properties on the spot: pride in one's detailed knowledge of squid biology, pictures of squids in conspicuous places, interest in squid-related internet news, passionate feelings about squid-octopus comparisons, etc. Some stereotypical properties might not be linguistically endorsed in the abstract, but rather only revealed by intuitive judgments about diagnostic cases – one use of clever thought experiments.

Whose hypothetical judgment about the properties characteristic of X constitutes the stereotype for X? For purposes of this paper, I'll treat as the default our well-educated English-speaking peers. However, for novel or half-novel or sub-culturally local property types (“alief” per Gendler 2008a-b; “narcissistic”, when the term was first introduced; being a hard-nosed reductive materialist) a narrower peer group will often be appropriate.

Just as one might have a disposition despite failing to manifest it due to a countervailing force or failure of normality, one might excusably fail to possess some of the dispositions in a stereotype. There is no principled ontological divide between dispositions broadly specified and more narrowly specified dispositions gathered together in clusters. *Ceteris paribus* defeasibility of a portion of a broad disposition, then, can be tantamount to *ceteris paribus* defeasibility of the entirety of a narrowly specified disposition within a larger cluster.

Depending on our interests and values, we might, in attitude ascription, choose to emphasize one aspect of a stereotype relatively more than another. For example, we might be more concerned about a person's patterns of explicit endorsement than about the person's in-the-world lived behavior or vice versa. See Section 8 below on intellectualism.

4.3. Further attitude-specific conditions. Sometimes having an attitude seems to require meeting non-dispositional conditions also. For example, presumably one can't *know that P* unless P is true. Maybe one can't have *marital love for Benjy* unless Benjy is alive. Maybe one can't *believe that the ocean is made of water* (as opposed to superficially similar "twater"; Putnam 1975) unless one has had causal contact with H₂O. I won't attempt a systematic analysis of such additional conditions.

However, it is often possible to capture apparently non-dispositional conditions in dispositional language. For example, it's not clear that one can be disposed to make future plans with Benjy if Benjy is dead – one might be disposed to *try* to make future plans with Benjy, but that's a different matter – and the disposition to make joint plans might be essential to marital love. Perhaps *knowing some fact P* involves the capacity to act in a way that reflects sensitivity to the truth of P, a disposition one can't have unless P is true. Maybe being disposed to say "the oceans contain water" with normal semantics requires having had causal contact with H₂O, such that without having had the right causal contact one cannot fully satisfy the linguistic aspects of the dispositional stereotype for the belief that the oceans contain water.¹⁰ So dispositional properties can perhaps carry more of a load than one would think.

5. *Comparison to Personality Traits.*

¹⁰ See Schwitzgebel 2002 for more discussion of the H₂O case.

The present view treats having an attitude as structurally similar to having a personality trait.

To be extraverted, plausibly, is nothing more or less than for one's dispositional profile to match, sufficiently well, the dispositional stereotype for extraversion. Stereotypically, to be extraverted is to enjoy meeting new people, to enjoy parties, to be talkative, and to take the lead in social situations; it is to be uninhibited in expressing one's feelings and to tend to plan ways to bring people together; it is to tend to choose certain types of work and play over others; and so on (all *ceteris paribus*, of course). Similarly for being courageous or curmudgeonly or Machiavellian or mellow. Having such personality traits is just a matter of matching the characteristic dispositional stereotypes well enough, or so it seems plausible to say.

Compare, now, the attitude of believing that one is God's gift to women. A man who believes this might not be disposed to say to himself, "I am God's gift to women", but he will presumably take for granted his attractiveness to women. He will tend to hold himself with a certain sexual arrogance. He will expect his advances to be favorably received. When his advances are rebuffed he will be prone probably not to surprise (which might render the dispositional structure unstable by correcting him in the long run) but rather to rationalization or quick forgetting. Note that for this attitude, unlike most beliefs, sincere assertion is not central to the stereotype. Note also that the category label employs somewhat fanciful language. The man might be an atheist who in some sense could not literally believe he is God's gift to anyone – and yet there is some belief-like attitude attributed through this fanciful language, an attitude that is probably not precisely captured by any more literal-seeming attributions like "he believes that most women are attracted to him" or "he believes that women are lucky to receive his sexual attention". My thought is that to attribute this belief is to gesture toward a dispositional portrait

the central features of which are broadly recognizable to normal attitude ascribers – much as to say that a person is extraverted or Machiavellian is to gesture toward a recognizable dispositional stereotype. This particular case reveals, I think, one advantage of the posture metaphor over the Belief Box metaphor. It seems reasonable to say that our imagined man has the attitude or profile or psychological posture characteristic of believing one is God’s gift to women; but it seems misleading to say that a representation with the propositional content “I am God’s gift to women” is written in his Belief Box, ready to be accessed and deployed in theoretical and practical inference.

Believing one is God’s gift to women might be a borderline case between having an attitude and having a personality trait. The similarity between personality traits and attitudes also appears in other seemingly intermediate cases: wanting constant attention, liking to take it easy, loving children, being unconcerned about the future, being self-confident. Are these broad-reaching attitudes or narrow personality traits? If attitudes and personality traits are structurally similar, we might expect such borderline cases. If attitudes and personality traits are entirely different psychological structures, these seemingly borderline cases create classification problems.

6. Mad Attitudes, Alien Attitudes.

I am recommending an approach to the attitudes that embraces a broad notion of the surface phenomena and then rejects the impulse to go deeper – or rather, as I will explain in the conclusion, rejects that impulse for present purposes. On the “surface”, as I define it, are all the stereotypically recognizable features of having an attitude, especially behavioral dispositions, experiential or “phenomenal” dispositions, and dispositions to enter other stereotypically

recognizably related psychological states, including acquiring or manifesting related attitudes – what I will call “cognitive” dispositions.

As mentioned in Section 2, deep and superficial accounts will disagree about classification when deep properties and surface properties diverge. Now, might deep properties and surface properties diverge *radically*? Let’s consider two diagnostic cases.

Case one: Mad belief. Andi, let’s suppose, is in some *Deep Condition D* that is held by a proponent of a deep account to be sufficient for believing that giraffes are born six feet tall. Perhaps a 22nd century brain scanner has found in Andi’s Belief Box a sentence, in the Language of Thought, translatable into English as “giraffes are born six feet tall”. Or maybe some neural condition is satisfied – some complex and subtle version of the giraffe neuron being linked to the six-feet-tall neuron being linked to the birth-size neuron. Now, *if* it is possible – at least conceptually possible – for deep and surface properties to diverge radically, then we should also be able to suppose that, despite satisfying this sufficient deep condition for belief, Andi is not at all inclined to act and react in the usual way. She is not at all disposed, for example, to say that baby giraffes are six feet tall. If asked explicitly, she would say giraffes are probably born no more than three feet tall. If shown a picture of a giraffe as tall as an ordinary man she would assume it’s not a newborn. If a zookeeper were to tell Andi that giraffes are born six feet tall, Andi would feel surprised and would say, “Really? I would have thought they were born much smaller than that!” And so forth, robustly, across a wide range of actual and counterfactual circumstances. None of these facts about Andi are due to the presence of *ceteris paribus*

defeaters like guns to her head or manipulation by evil neuroscientists or a bizarre network of other attitudes like thinking that “three” means six.¹¹

Case two: BetaHydrian valuing. Tomorrow, aliens from Beta Hydri arrive. The BetaHydrians show all signs of valuing molybdenum over gold. They will trade two ounces of gold for an ounce of molybdenum, with no apparent hesitation. When they list metal prices in their currency, they list the price of molybdenum higher than the price of gold. They learn English, and then they say things like “in BetaHydrian culture, molybdenum is more valuable than gold.” And so forth. Suppose, too, that BetaHydrians have conscious experiences. There is a kind of swelling they feel in their shoulders when they obtain things for which they have been striving. They translate this feeling into English as “the pleasure of success”. They experience this swelling feeling when they successfully trade away their gold for molybdenum. Like us, they have eyes sensitive to the visible spectrum, and like us they have visual imagery. They entertain visual imagery of returning to Beta Hydri loaded with molybdenum and of the accolades they will receive. Pleasurable feelings accompany such imagery. They plan ways to obtain molybdenum, at the cost of gold if that’s what it takes. They judge other BetaHydrians’ molybdenum-for-gold trades as wisely done. Etc. Ordinary people around Earth find it eminently natural to say that BetaHydrians value molybdenum over gold. But we know nothing yet about BetaHydrian biology or cognitive architecture, except that whatever it is can support this pattern of action, thought, and feeling. If Deep Condition E is some non-surface condition necessary for valuing molybdenum over gold on some deep account of the attitudes, and if we

¹¹ See Schwitzgebel 2012b for a similar example. “Mad belief” is built on analogy with “mad pain” in the sense of Lewis 1980. The Beta Hydri case is built on analogy with “Martian pain” in Lewis’s sense. Lewis (1994) seems to favor an approach to the attitudes similar to his approach to pain. If so, we agree about Martians but disagree about madmen.

can coherently conceive Deep Condition E's coming apart from the dispositional patterns above, then suppose Deep Condition E is not met. If we may conceive the existence of robust dispositional structures not underwritten by categorical bases, we might even imagine that the BetaHydrians robustly, intrinsically, durably, and non-accidentally exhibit these behavioral and cognitive and phenomenological patterns, across a wide range of possible worlds, despite being made entirely of undifferentiated balsa wood.

Ordinary opinion would, I think, favor saying the BetaHydrians value molybdenum over gold and favor denying the coherent conceivability of the Andi case: If Andi's dispositional structure is like *that*, she doesn't believe that giraffes are born six feet tall, whatever might be true about her deep structure. Furthermore, as folk psychologists ourselves, living in a social world, ascribing attitudes for the sake of praising, blaming, predicting, and explaining the kinds of things we care about, it seems like we *should* make those choices. There's a politics in deciding what in the world deserves such important labels as "desire", "value", "resent", and "believe". If superficial properties are what matter to us in ascribing attitudes, as I think they are, then if possible our classification decisions should track them.

Now, some deep accounts might disallow radical divergence between deep structure and superficial dispositions. Perhaps, for example, part of what it is for a representation with the content P to be in a Belief Box is for the subject to be disposed, *ceteris paribus*, to utter P when asked for her opinion on the topic. Faced with such an account, here's what I would do: Attempt to discover the maximum possible divergence the account allows between the deep structure and the superficial dispositional profile, and then consider whether, in such cases, dispositional profile or deep structure would produce better classificational practice given our interests.

Existing deep accounts, to the extent they commit clearly enough to permit such comparisons, will, I wager, tend to lose the contest.

7. *In-Between Cases.*

Alfred the half-grateful. Alfred is where he is today – a regional manager in a multinational company – partly through the help of Bertie, a somewhat older manager who had in his day been a regional manager in the same company. Bertie befriended Alfred early in Alfred’s career and bestowed on Alfred a variety of undeserved favors. After Bertie committed some embarrassing indiscretions unrelated to Alfred, Bertie was forced into early retirement. Bertie still sometimes reaches out to Alfred, wanting lunch, offering advice, seeking gossip. Alfred doesn’t quite resent such contact, but he agrees to it mostly from a sense of obligation. Alfred tells himself he is grateful for having had Bertie’s help. By dwelling on particular instances from the past, Alfred can sometimes bring himself to feel this gratitude vividly – a practice he has recently adopted before his occasional meetings with Bertie, though with only varying success. At other times, however, Alfred feels resentful and embarrassed by his connection to Bertie and wishes he had risen without Bertie’s help, which he is inclined to think he would have done. Alfred knows that he *should* feel grateful. He sincerely tells his spouse that he is quite grateful to Bertie. In fact, somehow, when he says it to his spouse it feels more true than when he says it to himself. Alfred has all these dispositions simultaneously. That is, he is presently disposed to act and react in characteristically grateful ways in some situations and some moods and, at the same time, disposed to act and react in characteristically ungrateful ways in other situations and other moods. Right now, let’s suppose, he is working on a client’s account and Bertie is far from his mind. Alfred is disposed to react ungratefully if Bertie were to

appear looking sleek and confident, and at the same time to react gratefully if Bertie were to appear looking rumpled and insecure. In sum, Alfred's dispositions do not align neatly into the stereotype. They splinter.

Juliet the implicit racist. Juliet is a philosophy professor who ethnically identifies as white. She is a political liberal, and she warmly espouses the view that people of all races are equally beautiful. She deplores what she regards as the racist beauty standards of Western society. When explicitly asked about the beauty of a black person, she will tend to avow the person's beauty – maybe even with considerable generosity toward people who are not conventionally beautiful – in a way she does not as reliably do for white people. She is more likely to spontaneously comment on the beauty of the black children of her few black acquaintances, e.g., in comments on pictures posted on social media, than to do so for her white acquaintances' white children. She sincerely professes to be awed by the beauty of certain black celebrities. And yet: It is the white celebrities whose beauty more viscerally engages her. When she goes to the beach, her swiftest, most spontaneous assessments favor the white bathers, though a deeply sown spontaneous aversion to any behavior that might be construed as revealing racist standards of beauty sometimes prevents her from voicing that admiration even to herself. She needs to work up her appreciation of the beauty of ancient, sun-wrinkled black men in a photography exhibit that is lovingly devoted to them. In contrast, she works to pull back her admiration of the chest-shaven white jeans models in department store advertisements. Presented the cases simultaneously, she would sincerely say she favors the hard-earned wrinkled

beauty over the vapid gloss of ads; but something else in her goes persistently the other way, and partly for racial reasons. In her most reflective moments, Juliet suspects all this about herself.¹²

Does Juliet find all the races equally beautiful? Is Alfred grateful to Bertie? The correct answer, I submit, is: *kind of*. Neither a simple “yes” or a simple “no” does justice to the facts. Compare, again, to personality traits. Few of us are 100% extravert or 100% introvert, 100% high-strung or 100% mellow. Rather, we match these profiles imperfectly and more closely in some respects than in others. If we match imperfectly enough, if we are stably prone to go sometimes one way, sometimes the other, often the best plan for describing us is to weasel out of any simple, overarching attribution and instead describe our patterns of splintering dispositions. There are gray, vague, in-betweenish cases. So also when our dispositions splinter away from neat alignment into attitudinal stereotypes.

The dispositional stereotype approach, partly because of its superficiality, handles in-between cases like Alfred’s and Juliet’s with a flexible minimalism: Display the dispositional structure and you’re done; nothing more to report! Deeper approaches, in contrast, invite the worry that something is still left open – for example, that underneath it all, Juliet might (or must?) really have “all the races are equally beautiful” in her Belief Box, or “white people are more beautiful” there, or maybe both; and until we have figured this out, we don’t know what her attitude really is, even if we perfectly well know her superficial dispositional structure. Furthermore, representationalist imagery – the Belief Box metaphor, the kinematics of belief P copulating with desire Q to beget intention R – tends to hide from view cases of splintering in-betweenness, by encouraging thoughts of discretely possessed representations with discrete

¹² I discuss similar cases of implicit bias in Schwitzgebel 2010, 2011, 2012c, and Myers-Schulz and Schwitzgebel forthcoming. Edited collections on the topic include Wittenbrink and Schwarz, eds., 2007 and Petty, Fabio, and Briñol, eds., 2009.

contents, retrievable from boxes, interacting as discrete units. In contrast, a very different style of thought flows from treating attitude possession in terms of dispositional profiles that imperfectly match folk psychological stereotypes – a style of thought on which in-betweenness is not anomalous and might even be our usual condition.¹³

8. *Three Types of Superficial Account.*

Let's say that a superficial account of an attitude is *behaviorist* (or near enough) if relevant surface phenomena are all (or mostly, or ultimately) cast in terms of outwardly observable behavior. And let's say that a superficial account is *intellectualist* (or near enough) if the relevant surface phenomena are all (or mostly, or ultimately) cast in terms of whether the person would endorse or utter or self-ascribe the attitude or its content. Finally, let's say that a superficial model is *liberal* if the relevant surface phenomena range broadly, including not only outward behavior and intellectual endorsements but, on a fully equal footing, a wide range of other phenomena. My own account is liberal in this sense. Behavioral and intellectual dispositions play a major role; but so also do phenomenal dispositions and cognitive dispositions other than explicit endorsement of the attitude in question.

Historically, superficial approaches to the attitudes have tended to be either behaviorist or intellectualist. Among the behaviorist ancestors are not only the self-avowed behaviorists, but also philosophers like Ryle (1949), Davidson (1984), and Dennett (1987), who continued to tend to privilege actual and hypothetical behavioral patterns as constitutive of attitudes. My view is

¹³ I extensively discuss in-between cases of belief in Schwitzgebel 1999, 2001, 2002, 2010; McGeer and Schwitzgebel 2006. See also Price 1969.

perhaps what follows from adopting a broadly Rylean-Dennettian approach, while treating inner phenomena as fully on a par with external behavior.

Both behaviorist and intellectualist accounts leave out too much of what properly concerns us in ascribing attitudes. We do and should care not only about outwardly observable behavior but also about private inner life, not only about patterns of intellectual assent but also about how one spontaneously lives one's way through the world.

Intellectualism is undesirable in another way too: A philosophical position that encourages us to think of ourselves as having the attitudes we intellectually endorse encourages noxiously comfortable self-portraits. If I steer myself through the world very differently than one might guess from what I (sincerely or for-all-I-can-tell sincerely) say to myself – if I treat women as stupid, enjoy my friends' failures, and repeatedly succumb to sexual temptation – on an intellectualist view I can still find solace in my high-minded egalitarian, magnanimous, and monogamous *attitudes*. The spirit was willing but the flesh was weak! Or: System 2 (my rational self) was willing but that darn uncontrollable (reflexive, appetitive) System 1 was weak!¹⁴ No, no. The flesh *is* the spirit, in large part, and the “rational self” partly a public relations device with itself as its own most gullible audience. What we say we should do is important, and so is what we feel bad about in retrospect after having already reaped the pleasures of vice, but to the people around us, our lived choices are typically more important; and we should reserve central words like “desire”, “prefer”, “value”, and “believe” to mark the most important and broad-reaching patterns. Broad-based dispositionalism thus encourages less pleasant but more salutary self-portraits: On the largest, most morally significant, life-permeating

¹⁴ On System 1 vs. System 2, see Frankish 2004; Evans 2008 – not that Frankish or Evans endorse the style of reasoning I criticize here.

issues, our overarching patterns, and thus I think we should say our attitudes, are rarely quite what we hope or think they are, and it can require rigorous self-examination and dedicated labor to pound them into shape.

8. *Overlapping Stereotypes.*

One remarkable feature of stereotypes is that they overlap. The same dispositions can belong to more than one stereotype. Consequently, by virtue of satisfying one stereotype, a person can, for free, nearly match the dispositional stereotype for a closely related attitude. I can't explore the consequences at length here, but let me mention three issues of interest.

Beliefs about the number of planets. I believe that there are eight planets. It seems that I also believe that there are fewer than 9 planets. But do I also believe that there are fewer than 10 planets? Fewer than 11? Fewer than 127? That there are $\sqrt{64}$ planets? More than just the four inner planets? That there are eight planets within the gravitational well of the nearest large hydrogen-fusing body? That there are eight known planet-like entities within half a light year? That Shakespeare probably had too low an estimate of the number of planets? This list is, of course, potentially infinite. Representation-in-the-box approaches seem committed to the view that there are determinate number of stored representations about the number of planets: maybe just one or a few language-like representations, from which the remaining propositions can be swiftly derived, generating a distinction between explicitly stored core beliefs and swiftly-derived implicit beliefs; or maybe instead a stored map-like representation that equally represents the number of planets as 8 and as fewer than 127, etc.¹⁵ Both views create difficulties: The just-

¹⁵ See Dennett 1978; Fodor 1987; Lewis 1994; Braddon-Mitchell and Jackson 1996; Blumson 2012; Carruthers this volume.

a-few view draws a sharp line across what seems to be a smooth gradation – a surprisingly occult line given its presumably fundamental cognitive or architectural importance. How are we to know whether “the solar system contains more than just the four inner planets” is one of the explicitly stored representations or merely swiftly derivable? And are we simply to take on faith that all the attitudes we properly ascribe are in fact either core or swiftly derivable from the core, or is there some way of testing this claim? The maps view, in contrast, seems to overgenerate beliefs: Presumably I don’t believe that there are $-i^2e^0\sqrt{64}$ planets – at least not until after I engage in some mathematical reasoning – though a map presumably represents that fact as much as it represents that there are 8. For related reasons it’s a challenge to represent inconsistent beliefs in map-like format.¹⁶

I recommend a different approach. The stereotype for believing that the number of planets is eight and the stereotype for believing that the number of planets is one less than nine presumably contain very similar clusters of dispositions, though with somewhat different centers and peripheries. Being disposed to look for the missing planet upon seeing a chart with seven but not upon seeing a chart with eight belongs perhaps equally to both stereotypes, while being disposed to affirm that the number of planets is one less than nine is central to one stereotype and at best peripheral to the other. The farther the belief ascriber’s locutions drift from what the subject would tend spontaneously to say and retrospectively agree to, the less neat the subject’s fit to the stereotype associated with those locutions and the less apt those locutions will usually be for sketching the subject’s psychological posture. Thus, the dispositional stereotype approach can respect the idea, which seems harder to accommodate on a stored-representation view, that what we have here really is just a smooth gradation from the apt to the awkward to the silly.

¹⁶ I explore these concerns slightly further in Schwitzgebel 2006/2010, 2012a.

Desiring, valuing, and believing good. Dispositional stereotypes can substantially overlap even if they seem to involve entirely different attitude types. Shortly after moving into one of my residences, I met a nineteen-year-old neighbor. Let's call him Ethan. In our first conversation, it came out (i.) that Ethan had a handsome, expensive new pickup truck, and (ii.) that he unfortunately had to go to community college because he couldn't afford to attend a four-year school. Although I didn't think to ask Ethan whether he thought owning a handsome pickup truck was more important than attending a four-year university, let's suppose that's how he lived his life in general. Ethan's inward and outward actions and reactions – perhaps not with perfect consistency – generally revealed a posture toward the world of valuing his truck over his education, or thinking that it's more important to have a beautiful truck than to go to a demanding university, or wanting a beautiful truck more than wanting to attend a four-year school. On a dispositional stereotype approach to the attitudes, we can treat the stereotypes associated with these somewhat different attitudes as largely overlapping, though with different centers and peripheries. Believing and desiring and valuing would seem on the surface to be very different attitude types, and are often treated as such – beliefs are “cognitive”, desires “conative”, they have different “directions of fit”, etc. – and yet in Ethan's case, the particular belief, desire, and valuation seem only subtly different.

Similar remarks apply to the BetaHydrians: Virtually the same clusters of dispositions, with subtle differences of emphasis, make it true that they value molybdenum over gold, that they regard molybdenum as more valuable than gold, that they desire molybdenum more than they desire gold. On a dispositional stereotype approach, there is no sharp division between these attitude types, though the attitudes seem to cross the cognitive-conative divide. Maybe an attribution of “desire” highlights a more visceral attraction while “valuing” and “regarding as

valuable” emphasize more intellectual aspects, though in slightly different ways. But if so, that’s a subtlety.¹⁷

Self-knowledge. Despite my pessimistic remarks about self-knowledge in Section 7, the fact that stereotypes can overlap has some positive epistemic consequences. The dispositions constitutive of believing that one believes that P and wants X and hates Y overlap substantially with the stereotypes constitutive of believing P, wanting, X, and hating Y, at least for normal adult human beings with the ordinary range of cognitive capacities. For example, stereotypical of both the lower-level attitudes and the higher-level beliefs about those attitudes are dispositions toward self-ascriptive utterances such as “I believe P”, “I want X”, and “I hate Y”. Thus, someone who fully matches the stereotype for lower-level believing, wanting, and hating will *ipso facto* already possess an important portion of the stereotype for the higher-order belief that she has those attitudes; and someone who is (*ceteris paribus*) disposed to refrain from the self-ascription *ipso facto* already partly deviates from the profile characteristic of the lower-order attitude. Let me emphasize, though, that for many of the most morally significant lower-order attitudes – such as how much one values one’s marriage – self-ascriptive dispositions are a minor part of the story, and so (in accord with Section 8) one can quite easily simultaneously match well enough the lower-order stereotype for having Attitude A and the higher-order stereotype for believing one has Attitude not-A.¹⁸

It is in general, I think, an appealing feature of the dispositional stereotype approach that, through the mechanism of overlapping stereotypes, it naturally handles the fact that possession of

¹⁷ This point seems relevant to debates about “motivational internalism”. See Bromwich 2010; Steinberg 2011.

¹⁸ See Schwitzgebel 2011 for a fuller discussion of self-knowledge of belief in particular.

one attitude seems nearly but not quite tantamount to the possession of related attitudes, both within and across general attitude types.

10. *Approximations.*

Despite my comments above about the number-of-planets case, sometimes circumstances call for the ascriber's phrasing things in ways that would not be endorsed by the subject herself. Ascription of attitudes to young children and infants, to non-human animals, to the conceptually or referentially confused, and across major linguistic barriers often fit this pattern: "Two-year-old Maya finds the new Latino teacher's fart jokes hilarious." We tolerate the fuzziness of the approximations because they are familiar and don't confuse anybody, and because we have no superior tools to work with, no alternative ascriptive technique that is more apt.

In certain moods, this can seem problematic. Does Confucius believe that *benevolence*, really, is present in him as soon as he desires it? The classical Chinese notion of *ren* isn't really quite the 21st-century English notion of benevolence. Does the dog really believe that a *bone* is buried by the *tree*, if his conceptualization of the world is so different from ours? Maybe, instead, the dog thinks a thing-that-smells-like-this is in the ground near the tall-leafy-pole-that-smells-like-that?¹⁹ Sometimes, I think, we react to such cases by bracketing certain parts of the stereotype. It might not be, so much, that dogs are *ceteris paribus excused* from linguistic and human-conceptual manifestations of their attitudes (although maybe we could run the case that way) as that we implicitly agree to disregard their deviation from that part of the stereotype; since no one thinks dogs can talk, no one is misled. We might even create animal-specific

¹⁹ See Stich 1979; Routley 1981; Davidson 1982; Smith 1982; Allen 1992; Andrews 2008/2011; Hutto this volume.

stereotypes, such as “territorial”, or modify human stereotypes for purposes of non-human attitude ascription.

Approximation isn't only a matter of cognitively or conceptually remote cases, though. Consider the case from Audi (this volume) of avoiding stepping on a rock when crossing a stream because the rock is too craggy. Audi holds, rightly I think, that there may be no determinately best choice between ascribing me the belief that the rock is craggy vs. the stone is craggy vs. the surface is craggy; there can be some inherent indefiniteness here, and maybe no such propositional ascription can be entirely accurate. I would add the further thought (which might or might not be consistent with Audi's overall view) that it can also be indefinite whether my attitude is that it is craggy vs. unsafe vs. slippery vs. a poor foothold. Is the thought, then, simply too thin to have specific propositional shape? Well, maybe, but attitudes can be also so rich and complex that they elude precise specification: It might be convenient shorthand to say that I believe of Girard that he is a buffoon; but really, my attitude toward Girard might be much more nuanced than that, too nuanced to capture in any brief way with existing vocabulary. “Buffoon” isn't really quite right, just the best I can do in one breath. Indeed, the human mind is so complex and unstable that maybe all our attitude ascriptions can only be imperfect approximating shorthand.

11. Occurrent Attitudes and Their Wraiths.

Philosophers sometimes distinguish *dispositional* attitudes from *occurrent* attitudes.²⁰ Dispositional attitudes typically endure over long periods and can be possessed even by people in dreamless sleep. We point to Lisa sleeping and say: “She thinks Lincoln was a great President”

²⁰ E.g., Ryle 1949; Price 1969; Audi 1994, this volume.

– a dispositional attitude ascription. Occurrent attitudes, in contrast, live only briefly, only as long as the topic is actively before the mind. When Lisa wakes, she might occurrently entertain the thought that Lincoln was a great President. Similarly, perhaps, one can dispositionally want to change professions, for years running, or occurrently have the urge to do so right now. One can dispositionally resent having been forced off the committee or occurrently be feeling resentful. If we accept this occurrent-dispositional distinction, it might seem natural, or even tautologous, to regard a dispositional approach to the attitudes as appropriate for dispositional attitudes but inappropriate for occurrent attitudes.

I would prefer to say: There are occurrences which are central manifestations of long-standing dispositional attitudes. And when those occurrences occur, we can sometimes attribute an “occurrent attitude”. But: such occurrences might not in fact align very well with one’s long-standing dispositions. So for clarity it might be best to use different words for occurrent vs. standing attitudes, such as “judgment” for the occurrent attitude and “belief” for the standing attitude. Furthermore, without a decent suite of at least short-term dispositions in place, even ascribing an occurrent attitude might be misleading.

Juliet, let’s suppose, is on the beach watching two young children play together, one black and one white. The words “black skin really is more beautiful than white” arise in her mind. For all she can tell, these words express a sincere and endorsed thought. I see three possible ways to flesh out the dispositional story: (1.) For a moment, most or all of Juliet’s dispositions align with the stereotype. She would, at that moment, viscerally find a black torso more handsome than an otherwise similar white torso. Her eye would not linger longer over the Swedish blonde bikini babe than over the blonde’s dark-skinned friend. Etc. It’s just that Juliet’s dispositions won’t stay that way long-term. (2.) In that moment, as has generally been

the case, Juliet's dispositions are mixed in the way described in Section 7. (3.) Requiring a revision of the original Juliet case: Although she is unaware of any insincerity in her thought, Juliet's dispositional profile, both now and durably, both viscerally and intellectually, is far from characteristic of someone who regards black skin as more beautiful than white. She has a few, perhaps momentary, dispositions – for example, her current disposition to utter aloud, with a feeling of sincerity, “black skin really is more beautiful than white” – but little else. Maybe she is reacting to some other beauty advantage that the black child has over the white one, such as symmetry, and misattributing the difference in her aesthetic judgment to the skin-color difference. Maybe in most judgments between matched individuals she would choose the white as the more beautiful; and maybe, at this very same moment, if her husband asked her if black skin was in general more beautiful than white skin, she would sincerely deny it, despite the thought or seeming-thought now running through her mind.

What should we say in cases (1), (2), and (3)? I propose:

In case (1), say Juliet momentarily or occurrently judges or thinks that black skin is more beautiful than white. Attitude ascriptions with more of a long-term feel about them – “Juliet finds all the races equally beautiful”, “Juliet believes that white skin is more beautiful” – we might still treat as in-between cases, as in Section 6.

In case (3), deny that Juliet has even the momentary occurrent attitude that comports with her seemingly sincere inner speech. What she has, instead, is what I will call a *wraith* of that attitude. She has some (maybe all, if that's possible) of the phenomenology or subjective experience characteristic of that attitude, and a wisp of the dispositional structure, but not enough to merit attitude ascription. Compare singing to oneself “I'm headed to Graceland, Graceland”

or a student's saying, with a feeling of approval but little understanding, a sentence from Kant. Compare saying to oneself, when worried, that everything will be fine.

Finally, case (2) would be an in-between case – a wraith half-full, maybe.

High-sounding clichés raise similar issues: “a human life is worth more than any sum of money”, “you can achieve anything you want to achieve”, “all men are created equal”. Often, it's not clear what such claims even mean. They might be wraiths or half-full judgments with an emotional flavor but without the dispositional traction of full-bodied judgment.

A man might profess love very sincerely, but very temporarily. The possibilities are similar to Juliet's. The dispositions might really be there, but only fleetingly – too fleetingly, perhaps to deserve a sturdy-sounding term like “love”. Alternatively, his words might be almost empty, possibly quite unknown to the man himself, a trick carnal lust has played on him, perhaps with his own implicit cooperation. Or it might be somewhere in-between. However it goes, his attitude, whether short-lived or long, is present exactly insofar as the relevant dispositional structure is present. Absent such structure, it is froth.

12. The Overthrow of Folk Psychology?

You might worry that I have fetishized folk psychology and abandoned empirical science. Timothy Schroeder (2004), criticizes “messy” theories of desire of the sort I favor because they seem to neglect the advantages of science in explaining how superficial dispositions hang together; and Peter Carruthers (this volume) raises a similar point against my previous work on belief. Even worse, perhaps, if folk opinion about the mind is a confused morass – as I'm inclined to think it is – then my approach would seem to inherit those same confusions.

The problem is, formal science, right now, does us no better. No brain imagery study yields a more useful set of categories for getting at what we care about in ascribing attitudes; nor does any current representationalist philosophical psychology, except as an optimistic promise or simplistic cartoon sketch of the mind. Only in the cases of “remembering” and “seeing”, perhaps, is empirical psychology mature enough to begin to threaten folk psychological patterns of classifying the attitudes. And in these cases, as I suspect we will discover generally, there is no one unified structure undergirding what’s picked out by our broad, folk psychological concepts, but rather a misaligned plurality. If the mind is a weird, kludgy chaos of dynamic agonisms and antagonisms, thought might not proceed via the manipulation of representations held in functionally discrete belief and desire boxes, and it might derange both folk psychology and empirical science to hastily assimilate the categories of one to the other.

Already now, though, science can legitimately lead us to adjust our superficial stereotypes, either by producing entirely new stereotypes or by modifying existing stereotypical structures to incorporate rising knowledge. Psychological research on sexism, for example, can coin a new type – “the implicit sexist” – and also modify our existing stereotypes of sexism and egalitarianism *simpliciter*. Folk psychological stereotypes won’t sit still, anyway, and are always to some extent influenced by scholarship and science, hence “phlegmatic”, “extravert”, “agnostic”, and our post-Freudian sense of how desires might manifest.

Modifications of folk psychology inevitably venture beyond mere cool description. Our folk categories are to an extent normatively self-fulfilling: Because we have them, we live into them, and for the most part we rightly feel we *should* live into them. We shape ourselves toward the stereotypical pattern of the baseball lover, the political liberal, the Wittgensteinian, the person who values higher education. By regulating ourselves accordingly, we become more predictable

and humanly comprehensible than we would otherwise be; and when we deviate sharply and unpredictably, we can be called upon to explain ourselves.²¹ To change our stereotypes is thus already to begin to change our norms – a possibility both hopeful and alarming, if future science ever encourages a radical overthrow of our folk categories.²²

²¹ See Brandom 1994; Morton 2003; Hutto 2004; McGeer 2007.

²² For helpful discussion, thanks to Robert Audi, Tim Bayne, Peter Carruthers, Rik Hine, Sean Kelly, Hichem Naar, Nikolaj Nottelmann, Tim Schroeder, and the various people who have commented on my related blog posts at The Splintered Mind.

References

- Allen, Colin (1992). Mental content. *British Journal for the Philosophy of Science*, 43, 537-553.
- Andrews, Kristin (2008/2011). Animal cognition. *Stanford Encyclopedia of Philosophy* (Summer 2011 Edition).
- Armstrong, D.M., C.B. Martin, and U.T. Place (1996). *Dispositions: A debate*. London: Routledge.
- Audi, Robert (1994). Dispositional beliefs and dispositions to believe. *Noûs*, 28, 419-434.
- Baier, Annette (1985). *Postures of the mind*. Minneapolis: University of Minnesota.
- Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge.
- Blumson, Ben (2012). Mental maps. *Philosophy & Phenomenological Research*, 85, 413-434.
- Braddon-Mitchell, David, and Frank Jackson (1996). *Philosophy of mind and cognition*. Cambridge, MA: Blackwell.
- Brandom, Robert B. (1994). *Making it explicit*. Cambridge, MA: Harvard.
- Bromwich, Danielle (2010). Clearing conceptual space for cognitivist motivational internalism. *Philosophical Studies*, 148, 343-367.
- Carruthers, Peter (2006). *The architecture of the mind*. Oxford: Oxford.
- Cartwright, Nancy (1983). *How the laws of physics lie*. Oxford: Oxford.
- Cummins, Robert (1996). *Representations, targets, and attitudes*. Cambridge, MA: MIT.
- Davidson, Donald (1982). Rational animals. *Dialectica*, 36, 317-327.
- Davidson, Donald (1984). *Inquiries into truth and interpretation*. Oxford: Oxford.
- Davidson, Donald (1987). Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, 61, 441-58.

- Dennett, Daniel C. (1978). *Brainstorms*. Cambridge, MA: MIT.
- 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.
- Dupré, John (2002). *Humans and other animals*. Oxford: Oxford.
- Evans, Jonathan St. B. T. (2008). Dual-process accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255-278.
- Fara, Michael (2005). Dispositions and habituals. *Noûs*, 39, 43-82.
- Fodor, Jerry A. (1987). *Psychosemantics*. Cambridge, MA: MIT.
- Frankish, Keith (2004). *Mind and supermind*. Cambridge: Cambridge.
- Gendler, Tamar Szabó (2008a). Alief and belief. *Journal of Philosophy*, 105, 634-663.
- Gendler, Tamar Szabó (2008b). Alief in action (and reaction). *Mind & Language*, 23, 552-585.
- Helm, Bennett (2005/2009). Love. *Stanford Encyclopedia of Philosophy* (Fall 2009 edition).
- Hutto, Daniel D. (2004). The limits of spectatorial folk psychology. *Mind & Language*, 19, 548-573.
- Lewis, David (1980). Mad pain and Martian pain. In *Readings in philosophy of psychology, vol. 1*, ed. N. Block. Cambridge, MA: Harvard.
- Lewis, David (1994). Reduction of mind. In *Companion to philosophy of mind*, ed. S. Guttenplan. Oxford: Blackwell.
- Locke, John (1690/1975). *Essay concerning human understanding*, ed. P.H. Nidditch. Oxford: Oxford.
- Margolis, Joseph (1973). *Knowledge and existence*. New York: Oxford.

- McGeer, Victoria (2007). The regulative dimension of folk psychology. In D.D. Hutto and M. Ratcliffe, eds., *Folk psychology re-assessed*. Dordrecht: Springer.
- McGeer, Victoria, and Eric Schwitzgebel (2006). Disorder in the representational warehouse. *Developmental Psychology*, 77, 1557-1562.
- Millikan, Ruth Garrett (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT.
- Millikan, Ruth Garrett (1996). On swampkinds. *Mind & Language*, 11, 103-117.
- Morton, Adam (2003). *The importance of being understood*. London: Routledge.
- Mumford, Stephen (1998). *Dispositions*. Oxford: Oxford.
- Murphy, Dominic (2006). *Psychiatry in the scientific image*. Cambridge, MA: MIT.
- Myers-Schulz, Blake, and Eric Schwitzgebel (forthcoming). Knowing that P without believing that P. *Noûs*.
- Naar, Hichem (forthcoming). A dispositional theory of love. *Pacific Philosophical Quarterly*.
- Neander, Karen (1996). Swampman meets swampcow. *Mind & Language*, 11, 118-129.
- Neisser, Ulric (1967). *Cognitive psychology*. East Norwalk, CT: Appleton-Century-Crofts.
- Petty, Richard E., Russell H. Fazio, and Pablo Briñol, eds. (2009) *Attitudes: Insights from the new implicit measures*. New York: Taylor and Francis.
- Price, H. H. (1969). *Belief*. London: Allen & Unwin.
- Prior, Elizabeth W. (1985). *Dispositions*. Atlantic Highlands, NJ: Aberdeen.
- Putnam, Hilary (1975). *Mind, language, and reality*. London: Cambridge.
- Roediger, Henry L. (1980). Memory metaphors in cognitive psychology. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 558-567.
- Ross, Lee, and Richard E. Nisbett (1991). *The person and the situation*. Philadelphia: Temple.

- Routley, Richard (1981). Alleged problems in attributing beliefs and intentionality to animals. *Inquiry*, 24, 385-417.
- Ryle, Gilbert (1949). *The concept of mind*. New York: Barnes & Noble.
- Schroeder, Timothy (2004). *Three faces of desire*. Oxford: Oxford.
- Schwitzgebel, Eric (1999). Gradual belief change in children. *Human Development*, 42, 283-296.
- Schwitzgebel, Eric (2001). In-between believing. *Philosophical Quarterly*, 51, 76-82.
- Schwitzgebel, Eric (2002). A phenomenal, dispositional account of belief. *Noûs*, 36, 249-275.
- Schwitzgebel, Eric (2006/2010). Belief. *Stanford Encyclopedia of Philosophy* (Winter 2011 edition).
- Schwitzgebel, Eric (2010). Acting contrary to our professed beliefs, or the gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly*, 91, 531-553.
- Schwitzgebel, Eric (2011). Knowing your own beliefs. *Canadian Journal of Philosophy*, supplementary volume 35 (*Belief and agency*, ed. D. Hunter), 41-62.
- Schwitzgebel, Eric (2012a). Do you have infinitely many beliefs about the number of planets? Blog post at The Splintered Mind: <http://schwitzsplinters.blogspot.com> (Oct. 17, 2012).
- Schwitzgebel, Eric (2012b). Mad belief? *NeuroEthics*, 5, 13-17.
- Schwitzgebel, Eric (2012c). Self-ignorance. In *Consciousness and the self*, ed. J. Liu and J. Perry. Cambridge: Cambridge.
- Smith, Peter (1982). On animal beliefs. *Southern Journal of Philosophy*, 20, 503-512.
- Stich, Stephen P. (1979). Do animals have beliefs? *Australasian Journal of Philosophy*, 57, 15-28.

Steinberg, Jesse (2011). Weak motivational internalism lite: Dispositions, moral judgments, and what we're motivated to do. *Canadian Journal of Philosophy, supplementary volume 35* (*Belief and agency*, ed. D. Hunter), 1-24.

Sutton, John (1998). *Philosophy and memory traces*. Cambridge: Cambridge University Press.

Wittenbrink, Bernd, and Norbert Schwarz, eds. (2007). *Implicit measures of attitudes*. New York: Guilford.