

Acting Contrary to Our Professed Beliefs

or

The Gulf Between Occurrent Judgment and Dispositional Belief

Eric Schwitzgebel
Department of Philosophy
University of California at Riverside
<http://faculty.ucr.edu/~eschwitz>
eschwitz at domain- ucr.edu
951 827 4288

January 27, 2010

Acting Contrary to Our Professed Beliefs

or

The Gulf Between Occurrent Judgment and Dispositional Belief

Abstract:

People often sincerely assert or judge one thing (for example, that all the races are intellectually equal) while at the same time being disposed to act in a way evidently quite contrary to the espoused attitude (for example, in a way that seems to suggest an implicit assumption of the intellectual superiority of their own race). Such cases should be regarded as “in-between” cases of believing, in which it’s neither quite right to ascribe the belief in question nor quite right to say that the person lacks the belief. The simplifications and assumptions inherent in belief ascription break down. Views of belief on which sincere avowal of a proposition is sufficient for believing it invite us to a noxiously comfortable view of ourselves, according to which once we have our avowals right we have right the cognitive condition, belief, which philosophers regard as central to philosophy of mind, epistemology, and philosophy of action.

Acting Contrary to Our Professed Beliefs

or

The Gulf Between Occurrent Judgment and Dispositional Belief

Often our words and actions don't mesh. We lie, of course, and we waffle – but that's just the start of it. With genuine conviction and complete sincerity you endorse some proposition *P*. Every time you think about *P*, you reaffirm it; to you, it seems unquestionably true. Yet if we look at the overall arc of your behavior – at your automatic and implicit reactions, at your decisions, at your spontaneous remarks on nearby topics – there's a decidedly un-*P*-ish cast. What should we say you believe in such cases? The question is central to our understanding of belief.

Section I: Examples.

Let's start with examples.

Juliet the implicit racist. Many Caucasians in academia profess that all races are of equal intelligence. Juliet, let's suppose, is one such person, a Caucasian-American philosophy professor. She has, perhaps, studied the matter more than most: She has critically examined the literature on racial differences in intelligence, and she finds the case for racial equality compelling. She is prepared to argue coherently, sincerely¹, and vehemently for equality of intelligence and has argued the point repeatedly in the past. Her egalitarianism in this matter coheres with her overarching liberal stance, according to which the sexes too possess equal intelligence and racial and sexual discrimination are odious. And yet Juliet is systematically

racist in most of her spontaneous reactions, in her unguarded behavior, and in her judgments about particular cases. When she gazes out on class the first day of each term, she can't help but think that some students look brighter than others – and to her, the black students never look bright. When a black student makes an insightful comment or submits an excellent essay, she feels more surprise than she would were a white or Asian student to do so, even though her black students make insightful comments and submit excellent essays at the same rate as do the others. This bias affects her grading and the way she guides class discussion. She is similarly biased against black non-students. When Juliet is on the hiring committee for a new office manager, it won't seem to her that the black applicants are the most intellectually capable, even if they are; or if she does become convinced of the intelligence of a black applicant, it will have taken more evidence than if the applicant had been white. When she converses with a custodian or cashier, she expects less wit if the person is black. And so on. Juliet could even be perfectly aware of these facts about herself; she could aspire to reform; self-deception could be largely absent. We can imagine that sometimes Juliet deliberately strives to overcome her bias in particular cases. She sometimes tries to interpret black students' comments especially generously. But it's impossible to constantly maintain such self-conscious vigilance, and of course patronizing condescension, which her well-intentioned efforts sometimes become, itself reflects apparent implicit assumptions about intelligence.²

Kaipeng the trembling Stoic. Having been won over by the Stoics, or by pessimists, or by believers in eternal glory – let's say by Stoics – Kaipeng quite sincerely judges, not just on one occasion but repeatedly, that death is not bad. He can recite arguments toward the conclusion, arguments he finds compelling and which he fully accepts as he makes them. Yet he trembles on the battlefield, and not just in anticipation of pain. He regrets the death of a good

person, and not entirely on behalf of those who have lost the benefit of her company. He takes measures to forestall his own death, and not wholly from a sense of duty. His actions and reactions for the most part are indistinguishable from the actions and reactions of someone who considers death bad. Like Juliet, Kaipeng could be perfectly well aware of these facts about himself and seek to change them – perhaps through meditative practices – yet they remain mostly unchanged.

Ben the forgetful driver. Ben reads an email saying that bridge he normally takes to work will be closed for a month. He immediately realizes that he will have to drive a different route. However, the next day he finds himself on the old route, heading toward that closed bridge. Of course, when he sees the bridge, he remembers the email. Maybe he smacks himself on the forehead at that moment. The day after, Ben does the very same thing again. If he were older, he might say he was having a “senior moment”, but he’s thirty. He also forgets to allow himself extra commute time. In lining up chores, he neglects to consider that his new route takes him right by the dry cleaners and he convinces his spouse to get the cleaning for him.

The recent literature on belief is full of cases of this sort, with various authors attempting to make sense of them in different ways, for example, Darrell Rowbottom (2007), Aaron Zimmerman (2007), Tamar Gendler (2008a&b), David Hunter, (2009), Fred Sommers (2009), and Brie Gertler (forthcoming). I will discuss those attempts shortly. I differ from all of these philosophers in regarding cases like these as vague, or (as I prefer to say) *in-between*, such that careful description of the subject’s mental state requires refraining from either ascribing or denying belief. Instead, one must articulate the subject’s dispositional structure in finer detail. Such a view both makes more theoretical sense than the alternatives and coheres better with our aims in belief ascription.

Section II: Theoretical Reasons to Accept the Possibility of In-Between Cases of Belief

Before addressing the particular examples at hand, let me offer some general theoretical considerations in favor of the view that “believe” is a vague predicate that admits of in-between cases – cases such that careful ascription requires refraining from simple attribution or denial of the belief in question. I favor a dispositional approach to belief (as articulated in Schwitzgebel 2002) so I’ll begin by making the case in those terms.

Dispositional accounts of belief – accounts of belief, that is, that equate believing with being disposed to act and react in various ways in various circumstances – are generally built upon a broad dispositional base. One way of articulating this is to say, in Gilbert Ryle’s terminology, that they’re not “single track” dispositions but rather multi-track – they “signify abilities, tendencies or pronenesses to do, not things of one unique kind, but things of lots of different kinds” (1949, p. 118). For example, in Ryle’s words,

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, p. 135; see also Price 1969; Schwitzgebel 2002).

Alternatively, one might think of belief as having a single dispositional track, but a very wide track. Ruth Marcus (1990) says that to believe that *P* is just to (in the right sorts of

circumstances) act as if *P* obtains; David Hunter (2009) says that it is to act and react in ways that would advance one's plans and achieve one's goals if *P* were the case. On either Ryle's view or the Marcus/Hunter view, we're faced with the question of what to say when a person appears to only partly possess the relevant dispositional structure – when he acts and reacts only in some respects as if *P* obtains or is prone to do only some of the things characteristic of belief that *P*.³

In some such cases, maybe, there are what we might call “excusing conditions”: Ryle's thin-ice believer doesn't warn the other skaters because he doesn't see the other skaters or because he'd enjoy watching them fall through (as in Rowbottom 2007). Dispositional claims, like most generalizations, hold only *ceteris paribus*, “all else being equal”, or only against a defeasible set of background assumptions. If he doesn't warn the other skaters out of schadenfreude or because he's blinded by the sun, that deviation from the typical dispositional manifestation counts not at all against ascribing him the belief that the ice is thin. But of course there must be limits to such excusers. Otherwise, we could save any generalization or dispositional ascription we wished simply by excusing every counterinstance. Articulating the principles underwriting the limits on excusing conditions is a difficult task, beyond the scope of this essay. (I'd say, roughly, that when a candidate excusing condition would undermine the potential usefulness of the generalization we should reject it as an excuser.) But the following is, I think, highly plausible: Given that there must be limits on excusers, it will not be an exceptionless law of nature that whenever a person has enough of the dispositional structure characteristic of the belief that *P* that the ascription of that belief to her is not out of the question, that person must thereby *fully* match the dispositional structure characteristic of *P* belief, except insofar as the *ceteris paribus* clause is legitimately sprung. In other words, there must be cases

in which the relevant dispositional structure is only partly possessed. There must, indeed, be something like a continuum between full possession of all the relevant dispositions and possession of none of them – with a multidimensional spectrum of cases between the two extremes. (This multidimensionality is one reason that simple numerical attributions of degree of belief or degree of confidence are often inadequate.) If, as also seems plausible, no bright line neatly partitions this multi-dimensional spectrum, “believes that *P*” is a vague predicate.

In non-belief cases where broad-track or multi-track dispositions are only partly possessed, careful descriptions tend to refrain from simple attribution or denial. My computer is almost always stable but tends to be unstable when I’ve recently run a print job. It is right to say, unqualifiedly, that it’s stable? Or that it’s not? Is it stable *simpliciter*? It may sometimes be appropriate to give a simple yes answer (my son wants to use it to play games on an airplane) or a simple no (a friend wants to borrow it to print her dissertation chapters), but the most careful and general description requires refraining from saying either that it’s stable or that it’s not, instead specifying the conditions under which it is and is not likely to crash, hang, garble a file, or develop a case of snails and consternation. Villia is scrupulously honest, except about her dating life. Is she honest *simpliciter*? Tim is stubborn on Tuesdays and compliant on Wednesdays. Is he stubborn full stop? Again, neither a simple yes nor a simple no seems quite right, though in certain circumstances a simple yes or no could suffice – similarly for simple, paradigmatic vague predicates like “tall”: Is a man tall if he’s five foot eleven inches? In some contexts a simple yes or no may suffice, but a more careful ascription will instead clarify with specific detail: “Well, he’s five-eleven”.

If to believe is to possess a multi-track disposition or a broad-track disposition or (as I myself prefer to put it) a cluster of dispositions (which can include cognitive and phenomenal

dispositions as well as behavioral ones), then there will be in-betweenish cases in which the relevant disposition or dispositions are only partly possessed. And if we treat such cases analogously to other cases of the partial possession of multi-track or broad-track dispositional structures, then we should say of such cases that it's not quite right, as a general matter, either to ascribe or to deny belief *simpliciter* – though (as in the other examples) certain limited conversational contexts may permit simple ascription or denial. Belief language starts to break down; the simplifications and assumptions inherent in it aren't entirely met; in characterizing the person's dispositional structure we may have to settle with lower levels of generality.

Not only dispositionalism about belief favors regarding “believes that *P*” as a vague predicate admitting of in-between cases. The same considerations apply to the other leading contemporary approaches to belief, including representationalism, functionalism, and interpretationism (for a review see Schwitzgebel 2006). The *functionalist* holds that to believe that *P* is to be in a state that occupies or is apt to occupy a certain causal-functional role – a state that is, for example (and over-simply), apt to be brought about by perceiving or hearing about or inferring that *P* and that is apt to lead to avowals of *P*, is apt to promote action *A* if it's discovered that *A* will achieve a desired goal if *P* is true, and that is apt to be combined with the belief *if P then Q* to conclude *Q* (for more sophisticated functionalist accounts see Armstrong 1973; Loar 1981). There will of course be cases of partial match in functional role: Some but not all aspects of the relevant functional role may be satisfied. If the splintering is severe enough and of the right character, such cases will be in-between. The *representationalist* holds that to believe *P* is to possess, in a belief-like way, an internal representational token (perhaps a sentence in the language of thought) with the content *P*. Representationalists are also generally functionalists in at least a broad sense of the term – a sense allowing for the relevance of the

evolutionary and developmental history of the state or type of state in question: What makes the internal representational state the belief state it is, is a combination of facts about how that particular state came about, the evolutionary or developmental history of states of that type in the organism or species, and what further cognitive relationships the state is apt to enter (e.g., Fodor 1968, 1987; Millikan 1984; Lycan 1986; Dretske 1988; Nichols and Stich 2003). Again it seems clear that the relevant roles may be only partly filled. (Unfortunately, the images and metaphors often invoked by representationalists don't sit easily with gradualism: Talk of inserting language-like representations into "belief boxes", "memory stores", "file folders", etc., suggests a yes-or-no architecture. But the invitation to a yes-or-no architecture is an infelicity of these metaphors, rather than a structural feature of the theories themselves.) The *interpretationist* holds that to believe that *P* is to exhibit or be prone to exhibit patterns of behavior that are usefully classified by the interpretive tools of folk psychology and in particular by the attribution of the belief that *P* (e.g., Davidson 1984; Dennett 1987). Clearly, again, one's patterns of behavior may match the appropriate patterns to a certain degree, folk psychological ascriptions may be more or less apt, we may fall into such interpretable patterns only partially (a point that Dennett himself emphasizes).

On all of the leading contemporary approaches to belief, it's natural to suppose that there will be a wide array of in-between cases where the dispositional or functional or functional-historical role is only partly filled, the relevant patterns of behavior, response, and cognition only partly possessed. Since I myself favor a dispositional approach to belief, I will adopt dispositional terminology for the remainder of this essay; but I believe the arguments can be straightforwardly adapted.

Section III: The In-Between Approach Versus Other Approaches.

The question, then, is whether the cases at hand – the cases of Juliet, Kaipeng, and Ben – are among the in-between cases. I suspect that many people, once they're sufficiently warm to the possibility of in-between cases of believing due to partial fulfillment of dispositional profile, functional role, or the like, will find it intuitively plausible to regard those cases as in-between.

Consider Juliet in particular. Does she believe that all the races are intellectually equal? She appears in some respects to have the appropriate dispositional structure: For example, she is disposed to judge or affirm, in good faith, both inwardly and outwardly, that all the races are intellectually equal and to react with shock and condemnation to expressions of blatantly racist opinion. In other key respects, however, she lacks an egalitarian dispositional structure, for example in many of her emotional reactions and spontaneous intuitive judgments about individual cases. On the face of it, the situation seems similar to that of my computer which is only pretty close to being accurately describable as stable or Villia who is only pretty close to being accurately describable as an honest person: It's an in-between case and a simple answer won't do. Or, better, a simple answer won't do when the ascriber aims to speak carefully and generally. In certain contexts – as with Villia and my computer and the tallness of a man who's five-eleven – a simple yes or no attribution may suffice. If we're just interested in what side Juliet will take in a debate, a simple "yes, she believes that all the races are intellectually equal" seems the right thing to say; if one black student is advising another about whether to take her class a simple "no, she doesn't think that black people are as smart as white people" seems a fair assessment. It appears to be a typical in-between case of a vague predicate. She doesn't fit

neatly into the yes or the no, so if we're concerned to describe her precisely, a yes or no won't do.

Although I find this view pretty natural and appealing, it must not seem natural and appealing to everyone, for not one of the various clever and insightful philosophers who have begun to write extensively about such cases in the last few years adopts this view. If the proposition that the subject verbally endorses but fails consistently to act in accord with is *P*, then the prominently discussed alternative views are:

- the *pro-judgment view*, on which the subject (determinately – and hereafter assume that unmarked applications of a predicate are determinate⁴) believes that *P* and fails to believe not-*P* (Zimmerman 2007; Gendler 2008a&b),
- the *anti-judgment view*, on which the subject fails to believe that *P* and instead believes not-*P* (Hunter 2009),
- the *shifting view*, on which the subject shifts between believing *P* and believing not-*P* (Rowbottom 2007), and
- the *contradictory belief view*, on which the subject believes both *P* and not-*P* (Gertler forthcoming; maybe Sommers 2009).

I will now address these views in turn. One quick caveat first, though: None of these philosophers addresses *exactly* the cases I've put forward, so there's room for any of them to deny that her preferred analysis transfers neatly from the specific cases she provides to all, or any, of my own specific examples. I will ignore this complication and treat each philosopher's analysis as generalizing to all the cases at hand. Of course, if any of these philosophers finds my approach preferable to theirs for some cases, so much the better.

Aaron Zimmerman (2007) and Tamar Gendler (2008a&b) both adopt the pro-judgment view – the view according to which Juliet, for example, determinately believes in the intellectual equality of the races and determinately fails to believe in the intellectual inferiority of black people. Zimmerman and Gendler adopt this view for similar reasons: Belief, they say, is, or should be thought of as, what governs or is reflected in our rational, thoughtful responses. Belief belongs to that part of our cognitive lives that is sensitive to evidence and argument. Something else, not belief – Gendler calls it *alief* – governs or is reflected in our mere habits, our automatic response patterns, our uncontrollable associations. When the habits and the automatic, arational response patterns come apart from the rational thoughtful responses, belief follows the latter. In Gendler’s terminology, the subject believes one thing – what she explicitly endorses – and simultaneously has a belief-discordant alief. Juliet has egalitarian beliefs and racist aliefs.

Zimmerman argues that this type of approach makes better sense of our “reactive attitudes” toward belief, that is, our tendencies to praise and criticize people for what they believe. In making this point, he emphasizes a Ben-like case – Ben, remember, is our absent-minded driver (in Zimmerman’s version the subject absent-mindedly goes to dispose of the trash in the place where the trashbasket was formerly kept). Zimmerman says that in such cases it would be inappropriate to belittle the subject as ignorant or to attempt to convince him through rational argument, and thus that the reactive attitudes appropriate to belief that not-*P* are absent. Zimmerman is surely right that attributions of ignorance and attempts at argumentative persuasion are mostly out of place in such cases – and the point generalizes to Juliet, and Kaipeng, and Zimmerman’s own Juliet-like example. Thus, *some* of the reactive attitudes appropriately directed to someone who believes not-*P* will be inappropriate in such cases. But not all of them. We can – and Ben himself can, in retrospect – criticize Ben for stupidly thinking

the bridge was still open. We can remind Ben of the evidence, which he is neglecting to consider, that the bridge is closed. We can criticize Juliet for being biased in her assumptions about the intelligence of black people. We might even hope for some success in interfering with a racist reaction of hers by presenting to her, at the right moment, a rational argument for the equality of the races. Kaipeng's fellow Stoics can scorn him for his apparently negative view of death, can charge him with a muddled or wavering attitude. Similarly, we can appropriately apply some but not all of the reactive attitudes rightly directed to someone who believes *P*, the endorsed proposition. We can praise Juliet for her acumen in assessing the empirical evidence, but we cannot praise her for having deep-set egalitarian convictions or, in practical as opposed to argumentative contexts, for her egalitarian attitudes about the races. Indeed, in some contexts it might be appropriate to criticize Juliet by saying that she doesn't *really* believe in the intellectual equality of the races or that she doesn't believe it "deep down". As predicted by the in-between model, the reactive attitudes split; they cut both ways – with some belief-that-*P* reactions appropriate and others not, some belief-that-not-*P* reactions appropriate and others not, perhaps splitting differently in different of the cases.

Gendler's main argument against treating habitual and automatic responses as central to belief is this: Beliefs, by their nature, are meant to track the truth and to change in response to evidence. Aliefs – that is, arational, automatic, or habitual response patterns – do not, she says, change in this way. They change in response to (though maybe she should say they are partly constituted by?) changes in habit (2008b, p. 566). In a Juliet-like case, the avowed racial attitude *is* responsive to evidence: She developed it, presumably, upon seeing the evidence, and she would change it if contrary evidence presented itself. In contrast, Juliet's unthinkingly racist responses are unaffected by the evidence against them.

This line of reasoning, it seems to me, considerably overdraws the distinction. Our habits, associations, and automatic responses *are*, to a substantial extent, responsive to evidence; and our verbal avowals or dispositions to judge are often *un*-responsive to evidence. When I start parking in a new lot, that precipitates a change in the spontaneous or habitual direction I begin to walk upon leaving the office. It may precipitate that change instantly, without the intervention of further conscious thought, or it may do so slowly. When I'm finally told that "LOL" abbreviates "laughing out loud" and not "lots of love", my spontaneous responses do adjust, either swiftly or slowly. Evidence, whether presented verbally or encountered directly in the world, shapes my habits and associations, typically though not always in ways that we would rationally endorse if we considered it explicitly. Presumably we reserve the word "habitual" for behaviors repeatedly executed and slow to change – but clearly not all automatic, associative responses are habitual in that sense. Presumably we reserve the word "arational" for responses that are not sensitive to evidence or argument in the right way, but then it begs the question to pair "arational" as consistently as Gendler does with "associative" and "automatic". Juliet's implicit racism is stubborn, but she is far less cognitively flexible than I imagine her to be if it wouldn't eventually change with enough exposure to sizzlingly smart black people. Ashby Plant and colleagues (2009), for example, claimed to find implicit prejudice against black people, as measured by the Implicit Association Test, to decline dramatically in the context of extended exposure to Barack Obama's Presidential campaign.

Likewise, we may be over-estimating Juliet's even-handedness if we think that her disposition to judge that the races are intellectually equal was developed primarily in response to, and that it's easily changeable by, the objective state of the evidence. As I imagine the case, that disposition is wrapped up in her overall politics and liberal values. Many liberals, I think, would

be very slow to change their judgments about the intellectual equality of the races were good evidence of Caucasian superiority to start coming in. Dispassionate rationality is not the rule in cognition about matters central to our conception of the world and ourselves. Part of the human condition, as opposed perhaps to the condition of angels, is that reasons and appeals to evidence are often only window-dressing for judgments arrived at through processes no different from the processes that shape our automatic and associative reactions. This is almost a truism in contemporary psychology, as suggested by the work of Nisbett and Wilson on self-knowledge of the processes driving our judgments (Nisbett and Wilson 1977; Nisbett and Ross 1980; Wilson 2002), the work of Damasio on the relationship between emotion and reason (Damasio 1994), and the work of Haidt and others on the processes driving moral judgments (Haidt 2001; Greene and Haidt 2002; Hauser 2006). This isn't necessarily bad news, given that our explicit reasoning is often deficient and our spontaneous impulses often wise. Gendler is, I'm sure, well aware of this psychological work; but it seems to problematize her contrast between rational and evidence-responsive beliefs on the one hand and arational, automatic, associative aliefs on the other.

The fundamental problem with the pro-judgment view is that it artificially hives off our rational and thoughtful responses from our habitual, automatic, and associative ones. In asserting that only the first are pertinent to belief, Zimmerman and Gendler attempt to separate what is really an inseparable mix. People judge in part automatically, associatively, and arationally, and they often show high intelligence in their habits and their unthinking, spontaneous responses. Words spring from our mouths seemingly unbidden, sometimes to our regret and sometimes reflecting a wit and intelligence that surprises even the speaker, as in the skillful flow of repartee. Without thinking explicitly about it, we attune ourselves in a variety of

ways to our social milieu and to the driving situation on the road. Teachers of formal logic and critical thinking swiftly and spontaneously recognize affirmations of the consequent in student homework, probably using pattern-recognition abilities not so different from those used to respond to other sorts of patterns; and in conversation, too, we detect fallacies and falsities using, it seems likely, a suite of automatic, habitual, and associative abilities, as well as more overtly rational ones. Furthermore, though Zimmerman and Gendler sometimes invite the impression that the arational, habitual dimension in cases of this sort – the not-*P*-ishness – pertains mainly to affect and outward behavior, we shouldn't forget that there are explicit judgments on *both* sides of the divide in such cases: specific judgments about individuals in Juliet's case – perhaps even fleeting explicitly racist thoughts which she quickly repudiates – judgments about the terribleness of a friend's death in Kaipeng's case, judgments about whether the drycleaners would be out of the way in Ben's case. Even if we can sometimes point to paradigmatic cases of rationality and its absence, there is no fundamental divide in the mind between the rational and the arational processes. It is artificial, then, to try to tie belief only to the former. It's more natural and it cuts the mind closer to its joints to hold that we believe that *P* if our actions and reactions *generally* reflect a *P*-ish take on the world, whether those actions and reactions are associative and automatic or deliberate and reflective.⁵

The anti-judgment interpretation of the cases at hand holds that the subject determinately believes the judgment she repudiates and fails to believe the judgment she endorses. Juliet, for example, believes that black people are intellectually inferior. I feel the appeal of this view. If forced to choose between attributing to Juliet the racist belief or the anti-racist one, I incline to attribute the racist one: It's what better captures the patterns in her daily behavior, which is presumably the main concern of the potential belief-ascribers interacting with her – her friends,

students, and co-workers. What does it matter, really, if she disavows racism in her more reflective moments? Reflective moments are rare. It's mostly just talk. She has a largely inegalitarian perspective on the world. Noting what's appealing in the anti-judgment view – even if we don't entirely accept it – helps reveal more of what's wrong in the pro-judgment view. If in cases like Juliet's we attribute belief primarily on the basis of avowals and explicit judgments, then our belief ascription no longer captures what it would seem the main purpose of belief ascription to capture: whether or not her general cognitive stance is racist or egalitarian. We leave out, it seems, what matters most in ascribing belief.

David Hunter (2009) embraces an anti-judgment view of at least some cases of the sort at hand – though he adds as a caveat, late in his analysis, that an in-between approach may also sometimes be appropriate. Hunter's central claim is that people can have what he calls “alienated” beliefs – beliefs the believer herself judges to be false or unjustified.⁶ Hunter holds that one can be alienated from one's beliefs much as, on Harry Frankfurt's (1976) view, one can be alienated from one's desires. Just as the smoker may desire a cigarette while also at a higher level repudiating that desire, the implicit racist may believe her race superior while at a higher level repudiating that belief.

One preliminary difficulty with the anti-judgment view is that it doesn't seem a very natural way to talk. No one says, “I believe that black people are intellectually inferior, but I repudiate that belief. I judge that belief to be false.” Kaipeng will not say, unless he is very much under the sway of an unusual philosophical theory, “I believe that death is bad, but I regard that as a false and unjustified belief.” Such statements seem almost “Moore-paradoxical”, like “It's raining but I don't believe that it is” (Moore 1952; Wittgenstein 1953). Of course, it might be responded that we *should* be willing to talk that way. Or it might be responded that the

unnaturalness of saying such things reflects pragmatic constraints on speech rather than the impossibility of the state of affairs described (as, indeed, may be true of the traditional Moore-paradoxical statements). Indeed, my own view suffers from linguistic unnaturalness in some of the same areas. Yet in Hunter's case I think this linguistic unnaturalness points to something more fundamental: The anti-judgment account omits a very important aspect of what it is to believe – the very aspect Zimmerman and Gendler emphasize. It omits what the subject explicitly endorses, how she is disposed to judge the overall state of affairs all things considered, what side she would take in an argument, how she is disposed to reason about the case in reflective moments, her best conscious assessment of the evidence. All these, furthermore, will generally be intertwined with daily behavior, even if not dependably; in real cases of the sort at hand, the divisions between “lower” and “higher” or daily behavior and mere talk are likely to prove messy, the cognitive patterns unstable, our attempts to clean it up with sharp distinctions likely to fail or to apply only to a minority of cases. Zimmerman and Gendler privilege the intellectual aspect of a person's psychology in belief ascription, while Hunter privileges the in-the-world spontaneous behavior. But I recommend that we treat both as an important part of what it is to believe. Shouldn't belief be seen as what animates my limbs *and* my mouth, what shows itself diversely in my action and my reasoning and my emotional responses, not just in some pried off subclass of these things?

Hunter explicitly grounds his anti-judgment position in a broad-track dispositional account of belief: To believe “is to treat the world as if it were the case that p. More fully, it is to be disposed to act and react in ways that would advance one's plans and achieve one's goals if p were the case” (2009, p. 38). But Juliet, Kaipeng, and Ben appear to be so disposed in some key ways, not in others. Hunter does not sufficiently address this issue. It seems, in fact, that any

case of the sort Hunter regards as “alienated” – any case in which one repudiates one’s own beliefs as false or unjustified – is, by virtue of that very fact, a case in which the repudiated belief is partly undermined. To repudiate *P* as false and unjustified entails not – at least not in the same breath – endorsing *P* as true and supported by the evidence, and thus it entails in at least that one important respect *not* being disposed to treat the world as if it were the case that *P*, and not to act and react in ways that would advance one’s plans and achieve one’s goals were *P* the case, assuming that embracing truths and avoiding falsehoods is among one’s goals. At one point, Hunter suggests that the distinction between alienated and endorsed belief crosscuts the distinction between fully believing and only half or in-between believing. I’d suggest, in contrast, that Hunter’s alienated beliefs are necessarily in-between; being alienated is one way of being in-between.

Do Juliet and company shift back and forth between belief in one proposition and belief in the other? Darrell Rowbottom (2007) endorses this treatment of at least some cases like those I have in mind, writing that “*it seems that a state of ‘in-between believing’ p might be described exactly as a state in which synchronic degree of belief in p is highly sensitive to changes in context*” (p. 134-135, italics in original). Ben slides from having a high degree of belief that the bridge is closed when he reads the email to having a low degree when he is driving toward the bridge. In his office, Kaipeng believes that death is not bad; on the battlefield, he loses confidence in this belief. Juliet believes that all the races are intellectually equal when engaged in public debate, much less so or not at all when surveying her students on the first day of class.

The main problem with this approach is that it leaves us without resources to describe the subject’s overall attitude. When Juliet is mowing the lawn, with general propositions about equality far from her mind, and neither seeing nor thinking about anyone of any skin color, what

does she believe – that the races are intellectually equal or that they're not? The shifting approach leaves us stymied. It doesn't seem right to say she has *no* belief about the races at such a moment: Beliefs are dispositional; one doesn't cease believing when one falls asleep or turns one's mind to another topic; we don't need to find out what someone is currently doing to say what she believes the way we need to find out what someone is currently doing to say whether she is jogging. We need to be able to speak about Juliet globally, not just about her shifting judgments or assumptions in particular individual conditions. What's her *general* attitude? Furthermore, it seems possible for Juliet in a single moment *both* to be having a racist reaction and to be sincerely judging that the races are intellectually equal – for example, when she's having a racist reaction and trying to suppress it or when she's grading a black student's essay on intellectual equality, undervaluing the essay but regarding its conclusions as true. This is a possibility the shifting model gives us no means to accommodate.

Should we ascribe to Juliet *both* the belief that the races are intellectually equal and the belief that they're not? This seems to be Gertler's (forthcoming) interpretation of examples of this sort – perhaps driven by the thought that having *either* the disposition sincerely to avow *P* or an overall structure of spontaneous reactions in accord with *P* is sufficient to underwrite belief.⁷ However, I see little to recommend this approach if it's taken naked: It invites only confusion to say simply that Juliet both believes that the races are intellectually equal and believes that they are not. For comprehensibility, we need to add qualifications: In such-and-such respects, Juliet acts and reacts as an egalitarian, in such-and-such respects she does not. This is the clearer answer to questions about what Juliet believes; it's also the in-between answer. Does it add anything of value – anything besides confusion – to append to this clear answer the claim that Juliet believes both *P* and its negation? I'm not sure I understand that claim any better than I

would understand, in the case of my conditionally reliable computer, a description of it as both reliable and unreliable. Could we say that *part* of Juliet believes one thing and part believes another? Sometimes people talk this way, but there are serious difficulties with taking such divisions literally, like: How do the different parts communicate? How much duplication is there in the attitudes held by the different parts and in the neural systems underlying those attitudes? Such questions are quicksand, especially for an a priori philosopher.

Let me add to this – just briefly since it risks begging the question – that on my own approach to belief baldly contradictory belief is in general impossible: One cannot simultaneously determinately believe that *P* and determinately believe that not-*P*. The dispositional properties characteristic of belief that *P* – for example the disposition to explicitly judge that *P* when the question arises, the disposition to treat the world as though *P* were the case – conflict with the dispositional properties characteristic of belief that not-*P*, such that no person can have a dispositional profile that simultaneously matches well both the profile of a *P*-believer and a not-*P*-believer.

Fred Sommers (2009) describes a range of cases of “dissonant belief” somewhat like the cases at hand – the atheist who prays in the foxhole, the theist whose theism does not penetrate his life – and diagnoses them as cases in which the subject believes one thing “*de mundo*” and another thing “*de dicto*”. Belief *de mundo*, per Sommers, involves taking the world to have some attribute in much the same way that a languageless animal might take the world to have that attribute. Belief *de dicto* involves taking some proposition to be true or false (and thus appears to require having and applying the concepts of truth and falsity). Thus, Sommers might say, Juliet believes *de mundo* that black people are intellectually inferior while also believing *de dicto* that the races are intellectually equal. Unfortunately, Sommers doesn’t make explicit what

he thinks someone with dissonant beliefs believes *simpliciter*. The two most obvious options are: She believes both *P* and not-*P* (or both that the world is *P*-ish and that it is not-*P*-ish), each in a different way, or that belief *simpliciter* cannot rightly be attributed and the cognitive situation must be described more finely. The first approach is a contradictory belief approach; the second is closer to my own.

Either way, however, Sommers's approach cuts things too neatly, at least for the cases at hand. Juliet does not only regard "all the races are intellectually equal" as a true proposition; she reacts viscerally and intuitively against blatant displays of racism. Ben, before he sets off to work, might be both disposed to take the correct route were the car in front of him to make the turn that Ben himself needs to take and, at the same time, disposed to head straight for the bridge were the car in front of him not to make the turn; he might be, dispositionally speaking, on the knife's edge, with both dispositions as *de mundo* as dispositions get. Atheists praying in foxholes, I imagine, typically accept or at least flirt with accepting the truth of the proposition that God exists; at least it seems that we can conceive of an avowed atheist's dispositions splintering that way – indeed in a whole variety of different ways. A person might regard a proposition as true but not act and react in the world as though it is true (except insofar as regarding a proposition as true is itself a kind of action or reaction). That's a nice thought, and some cases may fit it to considerable extent, but it can't work, I think, as a *general* model for dealing with cases of the sort at hand. There are, in my view, many ways in which one's dispositions might splinter so as to leave one between believing and failing to believe. Some of the more common ways of splintering deserve names or conceptualizations – self-deception, unconscious belief, low confidence, mood-dependent acceptance, knowing how without knowing that, believing *de re* but not *de dicto* (see Schwitzgebel 2002 for more discussion) – and perhaps

Sommers's *de mundo / de dicto* distinction deserves a place among them. But the problem is much more general and most cases fit only roughly if at all.

I suspect that part of the impulse behind most or all of these views is the desire to handle *neatly* cases like the ones at hand, the desire to support a simple yes-or-no (or numerical degree) answer to questions about what these subjects believe. Maybe it doesn't seem satisfying, or like we've given a full analysis, if we can't say in the end "yes, she believes" or "no, she doesn't". But not all questions deserve simple yes-or-no answers. In cases like Juliet's, and Kaipeng's, and Ben's, we should resist the temptation to make punctate yes-or-no attributions. While they do not fully and determinately believe, neither do they simply fail to believe. They are somewhere in between. Their dispositions are divided and our attribution must be nuanced. Why should we expect, anyway, simple, punctate models of cognition always to work smoothly, to be anything other than fallible simplifications of the richly complex structure of the mind?⁸

Section IV: Pragmatic Considerations.

It's not a passion for factually correct metaphysics that animates me here. Actually, I think there's no such thing as factually correct metaphysics. There are just better and worse ways of conceptualizing the world, given our values and the empirical facts. I hope the interest of this essay doesn't hinge on that controversial metaphilosophy; but you may not fully understand the argument now coming if you don't see that it is intended as a *pragmatic* argument. This argument is directed primarily against the pro-judgment view, which is, it seems to me, both the most intuitively appealing and the most problematic of the alternative views.

The central empirical fact is this: A person may be absolutely, completely persuaded of the truth of a proposition, in the sense of reaching a sincere, unequivocal, unmitigated, unqualified, unhesitant judgment, and yet that judgment may fail to penetrate her entire dispositional structure. One may find oneself, against one's will, unable to shuck old habits of thinking and reacting. One may even recognize in advance that these habits will persist, despite one's current sincere judgment which rationally requires their alteration. This is what's going on in the cases of Juliet, Kaipeng, and Ben, and many of the other authors' cases too. Thought about such cases reaches back through philosophical history to Plato, Aristotle, Xunzi, the Stoics, Descartes, Hume, and many others, where it's often conceptualized as the conflict of reason and the passions. But I follow Antonio Damasio (1994), Jesse Prinz (2004), and others in thinking that the distinction between reason and the passions has been considerably overdrawn; we reason partly through our passions; our passions are implementations of our reason.

The practical question is this: Do we want to highlight this empirical fact about ourselves – what I'd call the gulf between occurrent judgment and dispositional belief – or do we want to marginalize it as anomalous? Given the importance of the term “belief” to philosophy – the term is absolutely central to philosophy of mind, epistemology, and philosophy of action – the terminological decision to classify these cases as cases of belief will, I think, inevitably marginalize them. It suggests, implicitly, that cases like Juliet, Kaipeng, and Ben can be discussed in the same way, across philosophical subfields, as the more standard cases in which judgment and other dispositions are neatly aligned. It assimilates the one to the other – and as in most cases of assimilation, we should expect the minority party to become an afterthought. Although Zimmerman and especially Gendler no doubt intend, in part, to emphasize the disconnection between judgment – they would say belief – and behavior, an avowal-oriented

approach to belief seems a poor way to go about doing so. If what we believe is what we avow, and what we believe is also this central topic of philosophy of mind, epistemology, and philosophy of action, then I worry that we are invited to a noxiously comfortable view of ourselves: Once we have our judgments right, we have our beliefs right, and thus we have right that aspect of our minds about which the philosophical community cares. The dispute is in part terminological, but at stake is one of the most important pieces of terminology in all of philosophy.

I recommend that we resist taking for granted – as I think we often do – that dispositional belief flows effortlessly from occurrent judgment. Part of my resistance involves insisting upon different names for the occurrent episode (the judgment, that is, the passing affirmation or taking-as-true) and the standing dispositional state (the belief). It is salutary, I suggest, to find it somewhat amazing that for so many of our judgments our dispositions do fall into line, all in a twinkling so to speak, so that belief follows immediately. Someone tells me that John has replaced Georgia as department chair. Swiftly, a whole fleet of dispositions change: I'll go to John not Georgia if I have a question about my merit file, I'll put forms for signature in his box not Georgia's, I'll go down to the chair's office to look for him, etc. Ah, the awesome power of our brains! (Still, didn't I just yesterday find myself heading for the chair's office to find Georgia...?) Our morally most important beliefs, however, the ones that reflect our values, our commitments, our enduring ways of viewing the world – they're not like this. They change slowly, painfully, effortfully. You say to yourself that your marriage is worth the effort of preserving, that you believe in God, that your students deserve your respect, and you mean it. But it takes *work* to bring one's overall dispositional structure in line with one's broad, life-involving judgments. Unless we do that work and bring about that change, people rightly rebuke

us for not really or fully or deeply believing what we say we believe. As Robert Brandom (1994), Richard Moran (2001), and Victoria McGeer and Philip Pettit (2002) have emphasized, avowal is *commissive*, entailing a certain amount of forward-looking self-regulation. It's partly prospective, something a speaker must work to make true, contingent in part on the speaker's ongoing commitment to corral contrary inclinations, using perhaps a wide variety of self-regulative strategies. This commissive, prospective element can tinge the avowals with anxiety: To make them true, you have to *live up to them*.

I say I value family over work. When I stop and think about it, it seems to me vastly more important to be a good father than to craft a few more essays like this one. Yet I'm off to work early, I come home late. I take family vacations and my mind is wandering in the philosopher's ether. I'm more elated by my rising prestige than by my son's successes in school. My wife rightly scolds me: Do I really believe that family is more important? If I knew myself better, I'd have to describe myself as torn and in-betweenish. Or: I sincerely say that those lower than me in social status still deserve my respect; but do I believe this if I don't live that way?⁹ Do I really have right the cognitive attitude I should care about most? The attitude that philosophers regard as central to epistemology and moral psychology? The best answer seems to be: kind of.

Why *do* we care what we believe? Why should the concept *belief* play such a central role in philosophy? Beliefs, as Ramsey (1931) said, are the maps by which we steer. We want to get things right, so we can maneuver successfully through the world. Verbal and intellectual maneuvering is one important sort of maneuvering, but so also is non-verbal and less overtly intellectual maneuvering; and if we settle ourselves comfortably with the thought that we *believe* correctly when we are merely disposed to *judge* correctly, we might fail to realize that we're

often steering quite wrong – literally so in Ben’s case – so that our actions and reactions don’t reflect reality as we judge it to be. Also, even independently of the usefulness of our maps, we want to be people of a certain sort. We want to think and respond consistently and naturally in a way that reflects a certain view of the world. Endorsing that view, no matter how sincerely, does not prove its penetration. If the aim of attributing belief is to say something about how we steer through the world, then judgment cannot be sufficient for belief. I see no better way to highlight the gulf between judgment and belief than with a gradualist dispositionalism that rejects sharp lines and encourages explicit discussion of our divergent and splintering responses.¹⁰

References:

- Armstrong, D.M. (1973). Belief, Truth, and Knowledge, London: Cambridge.
- Brandom, Robert B. (1994). Making It Explicit, Cambridge, MA: Harvard.
- Burge, Tyler (1979). Individualism and the Mental, Midwest Studies in Philosophy 4: 73-121.
- Carnap, Rudolf (1947/1956). Meaning and Necessity, Chicago: University of Chicago.
- Damasio, Antonio R. (1994). Descartes' Error, New York: Putnam's.
- Davidson, Donald (1984). Inquiries into Truth and Interpretation, Oxford: Clarendon.
- de Sousa, Ronald B. (1971). How to Give a Piece of Your Mind: Or, the Logic of Belief and Assent, Review of Metaphysics 25: 52-79.
- 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.
- Evans, Jonathan St. B. T. (2008). Dual Processing Accounts of Reasoning, Judgment, and Social Cognition, Annual Review of Psychology 59: 255-278.
- Fodor, Jerry A. (1968). Psychological explanation, New York: Random House.
- Fodor, Jerry A. (1987). Psychosemantics, Cambridge, MA: MIT.
- Frankfurt, Harry (1976). Identification and Externality, in The Identities of Persons, ed. A.O. Rorty, Berkeley, CA: University of California.
- 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 and Language 23: 552-585.

- Gertler, Brie (forthcoming). Self-Knowledge and the Transparency of Belief, in Self-Knowledge, ed. Anthony Hatzimoysis, Oxford: Oxford.
- Goldstein, Lawrence (1993). The Fallacy of the Simple Question, Analysis 53: 178-181.
- Greene, Joshua, and Jonathan Haidt (2002). How (and Where) Does Moral Judgment Work? Trends in Cognitive Sciences 6: 517-523.
- Haidt, Jonathan (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment, Psychological Review 108: 814-834.
- Hauser, Marc (2006). Moral Minds, New York: HarperCollins.
- Hunter, David (2009). Belief, Alienation, and Intention, unpublished manuscript.
- Keefe, Rosanna, and Peter Smith, eds. (1997). Vagueness, Cambridge, MA: MIT.
- Kripke, Saul A. (1979). A Puzzle about Belief, in Meaning and Use, ed. A. Margalit, Dordrecht: Reidel.
- Loar, Brian (1981). Mind and Meaning, Cambridge: Cambridge.
- Lycan, William G. (1986). Tacit Belief in Belief, ed. R. J. Bogdan, Oxford: Clarendon: 61-82.
- Marcus, Ruth (1990). Some Revisionary Proposals about Belief and Believing, Philosophy and Phenomenological Research 50: 132-153.
- McGeer, Victoria, and Philip Pettit (2002). The self-regulating mind, Language and Communication 22: 281-299.
- McGeer, Victoria, and Eric Schwitzgebel (2006). Disorder in the Representational Warehouse, Child Development 77: 1557-1562.
- Millikan, Ruth (1984). Language, Thoughts, and Other Biological Categories, Cambridge, MA: MIT.

- Moore, G.E. (1952). Reply to My Critics, in The Philosophy of G.E. Moore, ed. P. Schilpp, New York: Tudor.
- Moran, Richard (2001). Authority and Estrangement, Princeton: Princeton.
- Nichols, Shaun, and Stephen P. Stich (2003). Mindreading, Oxford: Clarendon.
- Nisbett, Richard E., and Lee Ross (1980). Human Inference, Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, Richard E., and Timothy D. Wilson (1977). Telling more than we can know: Verbal reports on mental processes, Psychological Review 84: 231-259.
- Nozick, Brian A., Anthony G. Greenwald, and Mahzarin R. Banaji (2007). The Implicit Association Test at Age 7: A Methodological and Conceptual Review, in Social Psychology and the Unconscious, ed. J.A. Bargh, New York: Psychology Press.
- Plant, E. Ashby, et al. (2009). The Obama Effect: Decreasing Implicit Prejudice and Stereotyping, Journal of Experimental Social Psychology 45: 961-964.
- Preston, Ted M. (2005). Bridging the Gap Between Normative Beliefs and Actions. Ph.D. Dissertation, University of California at Riverside.
- Price, H.H. (1969). Belief, London: George Allen & Unwin.
- Prinz, Jesse J. (2004). Gut Reactions, Oxford: Oxford.
- Ramsey, Frank P. (1931). The Foundations of Mathematics, and Other Logical Essays, London: Routledge & Kegan Paul.
- Ridge, Michael (2006). Sincerity and Expressivism, Philosophical Studies 131: 487-510.
- Rowbottom, Darrell P. (2007). "In-Between Believing" and Degrees of Belief, Teorema 26: 131-137.
- Ryle, Gilbert (1949). The Concept of Mind, New York: Barnes & Noble.

- Schwitzgebel, Eric (1997). Words About Young Minds, Ph.D. Dissertation, U.C. Berkeley.
- 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). Belief. Stanford Encyclopedia of Philosophy, August 14, 2006.
- Searle, John R. (1969). Speech Acts, Cambridge: Cambridge.
- Sellars, Wilfrid (1969). Language as Thought and as Communication, Philosophy and Phenomenological Research 29: 506-527.
- Sommers, Fred (2009). Dissonant beliefs, Analysis 69: 267-274.
- Sorenson, Roy (1997/2006). Vagueness, Stanford Encyclopedia of Philosophy, February 8, 1997, rev. August 29, 2006.
- Steinberg, Jesse (2009). Two Solutions to Kripke's Puzzle, unpublished manuscript.
- Stroud, Sarah (2008). Weakness of Will, Stanford Encyclopedia of Philosophy, May 14, 2008.
- Wilson, Timothy D. (2002). Strangers to Ourselves, Cambridge, MA: Harvard.
- Wittgenstein, Ludwig (1953). Philosophical Investigations, trans. G.E.M. Ancombe, New York: Macmillan.
- Zimmerman, Aaron (2007). The Nature of Belief, Journal of Consciousness Studies 14 (no. 11): 61-82.

¹ Philosophers sometimes assert that believing that *P* is necessary for sincerely asserting that *P* (e.g., Searle 1969; see Ridge 2006 for a recent discussion). That view assumes the falsity of this article's thesis. Two alternative starting points for an account of sincerity might be: (a.) an approach according to which sincerity requires not belief but rather judgment (which is not – see below – sufficient for belief); or (b.) a dispositional approach according to which there's a cluster of dispositions that are sufficient for sincerity but not, without supplementation, sufficient for belief. For purposes of this essay, I recommend working with an unanalyzed intuitive view of sincerity, leaving open the question of its connection to belief.

² Juliet would presumably show impaired performance in classifying light-skinned faces with negative words and dark-skinned faces with positive words relative to the opposite pairing in the Implicit Association Tests of the sort that have recently received considerable attention in social psychology and also in recent work on belief (e.g., Zimmerman 2007; Gendler 2008b; Gertler forthcoming): See Nozick, Greenwald, and Banaji (2007) for a review. "Racist" performance on tests of this sort alone might not be enough to justify describing someone as racist in the sense that Juliet appears to be, but it might at least call into doubt the depth of one's egalitarianism. You can test yourself for implicit racist, sexist, ageist, etc., associations at <https://implicit.harvard.edu/implicit/demo>.

³ Some philosophers, however – e.g., Carnap (1947/1956), Sellars (1969), and de Sousa (1971) – express a temptation toward a narrower dispositional view, on which to believe is only to be disposed to avow. It's not clear, however, whether any of these philosophers completely yields to this temptation.

⁴ I am a logical pluralist who holds that we should choose logical structures suitable to the features of the domain of discourse, and I find it convenient here to drop two-valued logic and the law of the excluded middle. However, I have tried to phrase my points mostly in a way that doesn't depend on rejecting two-valued logic. I believe it is natural in borderline cases of the application of a vague predicate to say that it is not "determinately" the case that the predicate applies or that it fails to apply, and that it is "not quite right" to simply attribute or deny the predicate. If you too regard these as acceptable ways of talking and you insist upon a two-valued logic, presumably you can reconcile the one with the other or at least see how to translate my view into terms friendlier to your own. Here's one possible way to go: Treat only what I'm calling "determinate" cases of belief as cases in which it's true that the subject believes, and treat both the determinate cases of non-belief and the in-between cases as cases in which it's false that the subject believes. The resulting view would then be that in the in-between cases at hand, it is false that the subject believes *P* and false that the subject believes not-*P*, but true that she "in-between believes" *P* and also true that she "in-between believes" not-*P*. You won't long evade vagueness this way, however, since there will presumably be cases between determinately believing *P* and only in-between believing *P*. The phenomena don't admit of sharp lines (like many phenomena, actually), and when we're interested in the hazy areas where sharp lines break down, it's exactly the insistence upon those sharp lines that gets two-valued logic in trouble. See Keefe and Smith (eds. 1997) and Sorenson (1997/2006) for reviews of the logic of vagueness.

⁵ I would not deny that the man who stands trembling on a glass floor high in the air still believes that the glass floor is solid and can support him, to use one of Gendler's examples. The difference between this case and the cases I've presented is that in my cases there is a broad range of dispositions that deviate from the dispositional profile characteristic of the endorsed

judgment, while in the glass-floor case the deviation is narrow and arguably excused. Especially relevant here are cognitive dispositions – dispositions to make certain inferences and to reach certain related judgments (e.g., in Juliet’s case about the quality of an applicant or student) that show little parallel between the two types of case.

In some psychological domains, psychologists find evidence for the existence of two parallel cognitive processes, one fast, automatic, and non-conscious and one slow, deliberate, conscious – processes that then co-operate or compete (see Evans 2008 for a review). For example, in visual search, we might be both following a deliberate search strategy and simultaneously primed for a target to “pop out”. However, the plausibility of such models for some cognitive processes in no way undermines the general considerations I’ve raised in the body of the essay against the general division of the mind into an associative and a rational part; nor would such a division map neatly onto a pro-judgment view, since standard two-process models allow judgments to arise from either type of process. Keith Frankish (2004), who offers the most extended philosophical treatment of the division of human cognition into two strands of roughly this sort does not limit belief to the conscious, deliberative aspect. If anything, Frankish tilts the other way: What shows itself in our automatic actions and reactions, deriving from the kinds of structures we share with non-human animals, he calls “basic belief”; the attitude manifested in the more conscious, controlled aspect of cognition Frankish calls “superbelief”. Daniel Dennett’s (1978) view, which Frankish cites as a precedent, similarly distinguishes between belief, which very broadly informs our actions and reactions, and “opinion” which is primarily verbal.

⁶ Hunter is careful not to say that the subject *believes* her belief to be false. If she did believe that, rather than just regard, or judge, or assess it to be false, his view would risk

collapsing into the contradictory belief view. Believing that one's belief that P is a false belief seems to entail believing that P is false. And if believing that P is false is tantamount, as it seems to be, to believing not- P , then the alienated believer – at least the believer who regards her own belief as false and not merely unjustified – believes two contradictory things: P and not- P . As I'll discuss shortly, my own approach to belief renders baldly contradictory beliefs impossible, and I suspect that Hunter's approach does too, for similar reasons.

⁷ Gertler appears to express some ambivalence about this view. It may also be possible to interpret her as holding that in cases like Juliet's, the racist holds only the inegalitarian belief. In that case see my criticism of the anti-judgment view above.

⁸ A very different but influential literature in which it seems that philosophers fall into error through insisting on punctate yes-or-no answers to questions about belief is the discussion of Kripke's (1979) puzzle of Pierre: See Goldstein 1993; Schwitzgebel 1997; and Steinberg 2009.

Let me also briefly mention here two potential objections to my view, specific to the Ben case. First, it seems intuitive to say that Ben *knows* that the bridge is closed, even though on my view it's only an in-between case of believing – and there's a broad consensus in epistemology that knowledge that P entails belief that P . Perhaps this objection generalizes to Juliet, etc., but I think the intuition is clearest in the Ben case. One can imagine his wife saying to him, "Why aren't you taking the extra commute time into account? You *know* the bridge is closed." One might respond on my behalf with the thought that it's not clear that Ben does know – at least in his forgetful moments – that the bridge is closed. After all, it also seems natural for Ben to say of himself, once he realizes his mistake, that he forgot the bridge was closed, and forgetting something seems *prima facie* incompatible with knowing it. However, I myself prefer to say that

Ben does know, all the while, that the bridge is closed, even though he is only in an in-between state of believing. Contra most epistemologists, knowledge, on my view, does not require (determinate) belief: Knowledge involves a *capacity*, and belief a (modal) *tendency*. One can have the capacity without the tendency. (I owe this idea to David Hunter in conversation.) That my intuitions about such cases are not entirely anomalous is suggested by preliminary survey results I've been collecting with Blake Myers-Schultz.

Second, it might seem that I am committed to saying that people don't believe – or don't determinately believe – things that they're prone quickly to forget. Intuitively, Ben determinately believes that the bridge is closed the moment he reads about its closure and multiple times thereafter when he reminds himself – he just doesn't *retain* that belief. The shifting view is tempting, I think, for this case in particular. I can accommodate the intuition that Ben does determinately believe by noting that dispositions can be fleeting: For a moment, perhaps, all his dispositions are in line. They just don't stay that way. Once Ben's dispositions collapse back into an in-betweenish mix, his belief state is in-between. So one way of thinking about the Ben case is that he shifts back and forth between determinate and in-between states of belief. I'm not sure this is the best way to think about Ben, but it is a way my approach can make room for.

⁹ Case of this sort somewhat resemble cases of “weakness of the will” or *akrasia*, much discussed in the philosophical literature on moral psychology and freedom (for a review see Stroud 2008). While I doubt that all cases ordinarily described as weakness of will can be helpfully treated as cases where occurrent judgment diverges from an in-between dispositional belief, I suspect that some can be. One issue is the breadth of conditions under which one responds in the “weak” way, and the diversity of the weak responses. If the conditions are broad

and if the weak responses include implicit inferences and other judgments that seem to be driven by accepting as good or true what one outwardly disavows, such cases might helpfully be treated as in-between cases of believing.

¹⁰ My thoughts on this topic have been much informed by conversations with Tori McGeer and Ted Preston. The views expressed (committed to?) here are not, I hope, entirely unlike the views found in some of their work (esp. McGeer 1996; McGeer and Pettit 2002; Preston 2005), as well as that of H.H. Price (1969). Thanks also to helpful conversations with Jonathan Adler, Quassim Cassam, Naomi Eilan, Tamar Gendler, Brie Gertler, Peter Graham, Terry Horgan, Linus Huang, David Hunter, Krista Lawlor, Coleen Macnamara, Darrell Rowbottom, Susanna Schellenberg, Jesse Steinberg, Zoltan Szabo, Aaron Zimmerman, and readers of my blog, as well as audiences at the 2005 Canadian Philosophical Association meeting, Australian National University in 2008, and University of London, University of Leeds, and University of Warwick in 2010. I have explored related themes in Schwitzgebel (2001, 2002) and McGeer and Schwitzgebel (2006).