OVER THE PAST HALF DOZEN YEARS, there has been an increasing interest among philosophers in cases in which agents’ beliefs seem to be at odds with their behavior. These kinds of cases, which (following Albahari 2014) I will call discordancy cases, range from those familiar in everyday life (like phobic responses and, on some accounts of the phenomenon, self-deception), to those for which our evidence comes primarily from the laboratory (implicit bias; priming), to the genuinely pathological (clinical obsessive-compulsive disorder; monothematic and schizophrenic delusion). In these cases, agents may behave in ways that seem to conflict with their professed beliefs. For instance, the phobic person may flee from a spider she knows to be harmless, or the committed feminist may find herself preferring a male candidate over an equally qualified female due to implicit bias. In other cases the conflict is less stark, but there is nevertheless evidence for a belief-behavior mismatch. The self-deceived person may avoid situations in which her professed belief would be put to the test. The Capgras patient may insist that his wife has been replaced by an alien but make no move to report her to the police or to seek his real wife; the sufferer from anosognosia for hemiplegia may assert that her limb is fine but avoid tasks that require its use.

Discordancy cases have inspired a large literature, focusing, in particular, on self-deception, on delusion and (more recently) on implicit bias and the role of what Gendler (2008a; 2008b) has called “alief”. The aim has (usually) been to offer a general solution to the puzzle of discordancy cases: Gendler, for instance, appears to suggest that a mismatch between belief and alief explains these cases. In this paper, I have no such ambitions. Instead, I aim only to show that there is a class of discordancy cases with regard to which none of the existing accounts is adequate, and to offer a sketch of an explanation of these kinds of cases. I will suggest that the mechanism at work in this kind of case might also be at work in some other discordancy cases, but I will not attempt to assess how widespread its effects are.

The kind of case I will focus on has rarely or never (depending on how certain discussions in the literature are interpreted) previously been the subject of sustained philosophical reflection, to my
knowledge; nevertheless cases like this are very familiar (at least they are to me). Consider this paradigmatic case:

Anxious thought. Joe drives off from his house for a weekend away. After just a few minutes on the road, it occurs to him that he might have left the gas stove on. He’s never left the stove on before, but it’s possible (in some undemanding sense of ‘possible’: compossible not only with the laws of nature, but also with everything Joe believes and with ordinary experience) he could have: he used the stove this morning to make his breakfast, and it has a mechanism that requires it to be turned off manually. He has no memory of turning it off (or, perhaps, he can’t be sure that his memory of turning off the stove is a memory from this morning, rather than some other time). He knows it’s extremely unlikely that he left the stove on, but he can’t stop worrying about it.

Call the proposition that Joe has left the stove on proposition \( p \). Does Joe believe that \( p \)? He doesn’t take himself to believe it. He certainly isn’t prepared to assert it. But he can’t bring himself to ignore it. In past cases of this sort, Joe has sometimes (though not always) acted in ways that suggest he believes that there is some nonnegligible chance that the worry that eats away at him is well-founded: he has, for instance, returned to his house. If behavior is typically explained by postulating belief-desire pairs, we have some reason to attribute a belief to Joe that would rationalize his behavior, despite his protestations to the contrary.

Call cases like this neurotic anxiety (NA) cases. In this paper, I will argue that unlike the discordancy cases that have received most attention hitherto, NA cases are not explained by unconscious or exotic (in the sense of being unrecognized by folk psychology) mental states. Nor do they involve mental partition. Rather, the discordancy we see in NA cases is explained by nothing more esoteric than anxiety. The representation \( p \) is not poised to cause behavior by occupying the functional role of beliefs or belief-like states; rather it causes anxiety, which in turn causes behavior. The causal route from representation to behavior is deviant.

In the first part of this paper, I will argue that none of the more prominent accounts of discordancy cases can explain NA cases. These accounts entail predictions that are false in NA cases or entail that the agent is more, or less, rational than Joe actually is. Working through these accounts will yield important clues as to how representational contents like \( p \) are poised in these cases, and this, in turn, will motivate a fresh story about how this content causes behavior.

I.
The most obvious way of explaining (apparent) discordancy in cases like Anxious thought is in terms of degrees of belief. We can often explain behavior by reference to the rough probabilities that agents assign to propositions. An agent like Joe might assign a sufficiently high credence to the proposition that the stove is off to count as believing it, but given that the costs of being wrong are, potentially, very high (by his lights), his credence may not be high enough to justify his setting aside all doubts. The expected (dis)utility of being wrong, multiplied by the probability, might instead rationalize his returning to the house to check the stove. This kind of case — call it a credence case — is a discordancy case, because agents who feature in them act in ways that conflict with their explicit beliefs.

Credence cases surely occur. But Anxious thought is not a credence case. It is true that, given that being wrong about the proposition that the stove is off is potentially very costly, it is rational for Joe to treat it as a proposition that should not be relied on if his credence falls below a certain threshold. However, it seems to him (and to observers, should there be any) that his credence does not fall, and is not close to falling, below that threshold. The probability he assigns to the proposition that he left the stove on is negligible; far too small to justify his anxious rumination.
When we explain an agent’s unwillingness to rely on the truth of a proposition by reference to her credences and to the costs, by her lights, of being wrong, we rationalize her behavior. To that extent, credence-based explanations of discordancy cases actually explain them away: they show that there is no discordancy between the agents’ beliefs and their behavior. But Joe’s anxious rumination is not rational; both he and disinterested observers agree on that. One way to bring this home is to point out that there are cases in which the costs of being of wrong are as high or even higher yet in which he is disposed to rely on propositions despite the same (or an even lower) credence with regard to them. He might be anxious regarding the proposition that he has left the stove on, say — a proposition to which he assigns a very low probability — but not about the proposition that the brakes of the old car he is driving might fail, or the proposition that this old bridge will not support his weight, despite acknowledging that his credences in these latter propositions are higher than his credence in the former.

Eric Schwitzgebel’s (2001; 2010) account of discordancy cases is in some ways similar to a credence-based explanation and fails for similar reasons. On Schwitzgebel’s phenomenal dispositional account of belief (2002), to believe that \( p \) is to have enough of the dispositions that we associate with that belief: dispositions to reason, to assert, to act, to feel, and so on. Schwitzgebel suggests that agents in discordancy cases are “in-between believers”. The agents in the cases he examines have a raft of dispositions that support attribution to them of the belief they profess, and another set that support attribution to them of a conflicting belief. Julie professes a belief in racial equality and often acts in ways that are congruent with that belief (she is disposed to “call out” racist behavior and to vote for candidates that themselves oppose racism, say), but she also often acts in ways that are incongruent with her professed belief (she is surprised when a black student makes an excellent point and looks less favorably on black job applicants, for example). Kaipeng is convinced by Stoic arguments that death is not a harm, but he trembles at the thought of his own death and mourns the death of others (and not just because he will miss their company). Perhaps there are (many?) such cases, and perhaps Schwitzgebel’s account best explains them, but NA cases do not fit this model.

Obviously, agents like Joe have some dispositions which do not belong to the dispositional stereotype we associate with believing that \( p \). It is this fact that licenses calling the case a discordancy case. But Joe has very few dispositions that conflict with his professed belief, and very many dispositions that support it.1 In high-stakes cases (suppose he is running late for his flight, and he needs to be on the flight to make an on-campus interview for a precious tenure-track job), Joe will act in accordance with his belief. His betting behavior will accord with it. He will plan in ways that depend on its being true (he won’t plan to book accommodation for himself when he returns to his hometown; rather, he will intend to go back to a home that is not a smoking ruin).

Further, and I think importantly, I predict that under cognitive load, agents like Joe would act very differently from the actual agents who feature in the kinds of cases that Schwitzgebel aims to explain. Schwitzgebel’s cases turn on there being a discrepancy between dispositions to engage in assertion and related behaviors and other dispositions. Assertion is a controlled behavior; the disposition to assert supervenes on the set of mechanisms or processes that are often (somewhat misleadingly) identified with “System 2”: processes that are slow, effortful, typically conscious and logical (Evans 2003).2 Cognitive-load manipulations decrease the extent to which behavior is controlled, and increase the extent to which it is driven by “System 1”— automatic, effortless — processes. Under load, agents act in accordance with their

---

1. Scott-Kakures (2001) centres his discussion of twisted self-deception on several cases, one of which is very like Joe’s. Unlike Joe, however, the agent who features in it convinces herself that she has left the stove on. Scott-Kakures may take this kind of case to represent the bulk of cases of this sort. I am confident that’s false; I think that, in the great majority of these cases, we do not believe the proposition about which we nevertheless fret. Regardless, it is cases like this, and not like the ones Scott-Kakures describes, that are my target here.

2. “System 1”/“System 2” talk is misleading because the processes characteristic of each do not cluster together sufficiently reliably to warrant talk of systems; see Carruthers 2013.
prepotent dispositions. Schwitzgebel’s cases apparently feature agents who have some of the dispositions that support controlled processing in conflict with those that support more automatic processing: at least, this is true of real agents like Julie, the implicit racist, and Kaipeng, the trembling Stoic. But Joe will act in accordance with his professed belief even under cognitive load, indicating that the representation that makes Joe anxious is not poised to drive behavior in the way that System 1 representations are. Indeed, distract Joe by requiring him to multitask (a classic load manipulation) and you will probably bring it about that he ceases to worry, at least for a while.

The dispositions that conflict with Joe’s professed belief seem too flimsy to support attribution to him of an in-between belief that \( p \).

The same goes, a fortiori, for contradictory belief accounts (Gertler 2011). If we don’t have sufficient grounds for attributing a less-than-wholehearted belief that \( p \) to Joe, we certainly don’t have grounds for attributing not-\( p \) to him. But if Joe doesn’t believe that \( p \), perhaps he nevertheless alieves \( p \).

Aliefs are a kind of mental state hypothesized by Tamar Gendler (2008a; 2008b). Aliefs have representational content that is associatively linked to affective and behavioral dispositions. Gendler cites a number of examples, any of which will do to illustrate the notion. Consider the mental state of the person anxious about venturing out onto the glass-floored Grand Canyon Skywalk. This person is confident that the Skywalk is perfectly safe — perhaps as confident (or even more confident) as she is that her coffee is safe to drink. She believes that the Skywalk is safe, but she alieves \( p \) that it is not: she represents it as dangerous and has corresponding affective and behavioral responses. Gendler cites a number of other discordancy cases that fit the alief mould: a person may believe that the dog-feces-shaped fudge is safe and tasty, but find it difficult to put it into her mouth; she may know that no harm will come to her loved one from throwing darts at a picture of her, but find it difficult to do so; she may know that the drink is safe, but find it difficult to drink because of the label which identifies it as poison (or even as not poison); and so on: in each case, her alief has a representational-affective-behavioral content which is at odds with her beliefs and explains the discordancy.

It seems that Anxious thought fits this mould precisely. Joe doesn’t believe that \( p \); rather he is confident that not-\( p \). But he entertains a thought with representational, affective and behavioral content: that the stove is, or might be, on; fear or anxiety; the urge to return home to check. It is, we might suggest, the conflict between his belief and his alief that makes this a discordancy case.

Despite appearances, however, Joe’s anxious thought is not an alief. There are two main differences between NA cases and cases featuring aliefs. First, aliefs are evidence-insensitive: “Beliefs change in response to changes in evidence; aliefs change in response to changes in habit” (Gendler 2008b: 566). NA thoughts are not evidence-insensitive. Were Joe to return to check the stove, or were he to phone his partner and ask him to check, the evidence he is presented with might cause his anxious thought to vanish. Admittedly, NA thoughts are, by their nature, somewhat resistant to evidence (were this not the case, the evidence that justifies Joe’s judging that \( p \) is very unlikely would also assuage his anxiety). His thought may resurface on his driving away after checking. Nevertheless, NA thoughts may be very responsive to new evidence: seeing that the stove is off may cause the thought to vanish, and is very likely to give him some temporary relief. Whereas aliefs are supposed to be insensitive “to the possibility that appearances may misrepresent reality” (2008b: 570), NA thoughts are sensitive to the

---

3. I’m confident that Schwitzgebel would agree, given that he is willing to ascribe wholehearted belief to Gendler’s reluctant Skywalker, on the grounds that the Skywalker has too many of the dispositions associated with believing that the Skywalk is safe and too few of the contrary dispositions to justify withholding ascription of the belief (Schwitzgebel 2010, n. 5). Joe hews even closer to the dispositional stereotype than the Skywalker, so he’s an even clearer case of a wholehearted believer.

4. In using ‘alieve’ in this way, I follow Gendler herself. However, she cautions that this usage is loose and somewhat misleading: she adopts a four-place characterization of aliefs, according to which, strictly speaking, agents do not
appearance/reality distinction. The fear of the Skywalker may not decrease at all in response to new evidence, but Joe’s anxious thought will abate and may vanish given appropriate evidence.

We shouldn’t rest too much weight on this difference between aliefs and NA cases, since the former may not be as evidence-insensitive as Gendler thinks (Nagel 2012). It might be better to say that to the extent to which they have associative representational content, aliefs are evidence-insensitive. Gendler is agnostic on the nature of their content, but evidence-insensitivity seems to be the hallmark of associative content (associations alter in response to counterconditioning, not evidence, and update slowly, unlike beliefs; see Mandelbaum forthcoming). 5 NA thoughts don’t seem to be like that. When Joe sees that the stove is off, his anxious thought will immediately cease (at least for a time, and perhaps permanently).

More importantly, however, aliefs have associated action tendencies; they dispose agents toward particular responses, either as a result of the inculcation of habits or as the upshot of evolution (an alief is “an innate or habitual propensity to respond to an apparent stimulus in a particular way” [Gendler 2008b: 553]). The phobic person is disposed toward fear, given the right stimuli; the person high in disgust is disposed toward responding to something that looks like dog feces by flinching away from it. In every case, acting on an alief is acting in accord with a prepotent response, which (once activated) it takes effort to override. As we have already seen, though, NA cases are not like that: under load, agents like Joe do not behave in ways that belie their professed beliefs. That strongly suggests that Joe’s representation of p is not an alief.

Several promising accounts aimed principally at the discrepancy between the assertions of people who suffer from delusions and their behavior might also explain NA cases. In a recent book, Philip Gerrans (2014) has offered one such account and explicitly claimed that it might explain cases very much like Anxious thought. Delusions, Gerrans claims, are not beliefs; rather, they are default-network thoughts. The default network is the network of brain mechanisms involved in the production of narratives concerning the self. Delusions and allied thoughts arise when this network is able to spin a narrative unsupervised by (dorso-lateral prefrontal cortex [DLPFC]) mechanisms that normally have the role of testing thoughts for consistency and empirical adequacy and which thereby come to qualify representational states as beliefs. Because these mechanisms are hypoactive, or because the default-network thoughts attract disproportionate processing resources, the essentially imaginative products of the default network come to play an abnormal role in the person’s thought and behavior.

Gerrans is explicit in claiming that this kind of account can apply to cases that resemble NA cases. He compares the delusional patient to an anxious parent who is worried about the whereabouts of his child. The fact that she is out later than she promised generates anxious thoughts about what might have happened, but the parent doesn’t believe that his daughter has had an accident. Nevertheless, this highly salient — default-network — thought may motivate actions (ringing local hospitals, for instance) in much the same way as beliefs do. But there is a major problem with attempts to explain NA cases in this kind of way. If NA cases were explained as default-network thoughts, the agents who feature in them would be more like delusional patients than they actually are. Since the default-network thought monopolizes processing resources, and the thought is therefore not tested for empirical adequacy and consistency, the person shouldn’t grasp the extent to which their thought is irrational. They should be more like the OCD patient with poor insight than the classic OCD patient (the

5. Mandelbaum (2013) argues that Gendler must suppose that aliefs have associative content; otherwise we could explain discordancy cases by invoking nonconscious beliefs. But such associative contents won’t do, he claims, because explaining the kinds of cases Gendler cites requires us to postulate representations with syntactic structure. Mandelbaum ignores the possibility of some kind of mid-strength option, which would have sufficient structure to explain the phenomena he cites but not enough to underwrite inferential promiscuity. Elsewhere (Levy 2015), I have argued that such mid-strength representations actually explain the data he and Gendler cite. But Joe’s thought is too evidence-sensitive to be plausibly regarded as such a representation.
same problem applies, of course, to attempts to explain clinical OCD via Gerrans’ proposal: why do most patients have insight into the condition?) We ought to expect, at least, to see some wavering of insight or attempts at rationalization of the thought if NA thoughts are default-network thoughts unsupervised by the mechanisms that would transform them into beliefs.

Agents like Joe can, and do, think hard and rationally about their anxious thoughts and their likelihood. They thereby bring to bear the mechanisms that transform default-network thoughts into beliefs. These mechanisms are not hypoactive in agents like Joe: on the contrary, he deploys them intensively and effectively. It is, no doubt, this fact that explains why he has such good insight into the irrationality of his rumination. But this fact also entails that his thought is not a mere default-network thought as Gerrans conceives of it. There may be some discordancy cases beyond clinical delusions that fit his description; Gerrans’ case of the anxious parent might be one. But if it is, it is not an NA case after all. In NA cases, we see good insight combining with a thought that has the power to motivate behavior; in the kinds of cases Gerrans has in mind, we do not.

Currie (2000) offers another account aimed at explaining the discordancy seen in delusions. He argues that delusional patients suffer from an inability to monitor the source of their thoughts, and therefore come to mischaracterize their imaginings as beliefs. A patient may have a delusional belief, but the delusional belief is not (say) that my wife has been replaced by a replicant but rather that I believe that my wife has been replaced by a replicant. Imaginings, as Currie points out, can play some of the inferential roles of beliefs: children who imagine that a toy has had water poured on it go on to imagine that the toy is now wet, for example. So Joe’s imagining that p might motivate his anxious rumination or even his returning to check the stove.

There are obvious differences between the kinds of cases that Currie aims to explain and NA cases, which make it inappropriate to transfer the account from the first to the second unchanged. While a delusional patient might believe that he believes his delusion, there seems to be little reason to attribute to Joe the correlative higher-order belief about his thought; he doesn’t believe that he believes that p. Nevertheless, perhaps we can explain Joe’s behavior by reference to the inferential role that imaginings sometimes seem to play: perhaps Joe believes neither that p nor even that he believes that p; he merely imagines that p, and his so imagining explains his behavior (including his mental behavior). Since imaginings sometimes appear to play some of the inferential roles played by beliefs, we might have the makings of an explanation of his behavior even though Joe lacks the higher-order belief Currie attributes to the delusional patient.

The problem with this account, though, is that it is false that imaginings motivate behavior in the kind of circumstances in which Joe’s thought plays this role. To the degree to which someone is immersed in imagining something, their imagining takes on more and more of the roles that we associate with beliefs, such as driving inferences and behavior (Schellenberg 2013). They do not continue driving behavior—not, at any rate, by mimicking the inferential roles played by beliefs—when the person appraises them for truth in the sober light of day. Joe can scrutinize his anxious thought all he likes, soberly and conscientiously assessing its likelihood; he remains unable to shake it off. The same kind of scrutiny that ensures that Joe’s representation is not a default-network thought also ensures that it doesn’t play the kind of role in behavior that Currie attributes to delusions.

Similar remarks apply to attempts to model NA cases on Gendler’s (2007) account of self-deception. Gendler maintains that self-deception occurs when a person pretends that something is the case, and their pretense comes to play (much) the same role as is normally played by belief, in terms of both “introspective vivacity” and causation of action. But Joe’s thought has none of the marks of pretense. Gendler argues that the pretense that p motivates what she calls evidential override (244), where the subject actively avoids evidence against p or situations in which p is put to the test, or simply avoids thinking hard whether p. There is no analogue of any such behavior in Joe’s case. On the contrary, characteristically agents like Joe actively seek all the
evidence they can. They may engage in genuinely high-quality cognition aimed at assessing the proposition about which they are anxious, and thereby come to have a clear and distinct grasp of its negligible likelihood. Joe might be a paragon of rationality, so far as his attempts to assess his thought are concerned, yet continue to be haunted by it.

II.

NA cases can’t be explained using the resources of existing proposals designed to account for discordancy cases. Joe’s credence in \( p \) is too high to explain discordancy in that kind of way. Joe’s anxious thought is not an in-between belief or an alief, because it doesn’t drive behavior under cognitive load. Even if it is an imagining, or a default-network thought, it doesn’t drive behavior in the way that Currie or Gerrans envisage, because Joe subjects it to DLPFC supervision of a kind that is inimical to these states playing the kind of quasi-inferential roles they seem to have in mind.

That said, Joe’s representation — his thought with the content that I have left the stove on (or something in that ballpark) — might nevertheless be a default-network thought or an imagining. Whatever it is, it is not poised in the kind of way that such thoughts need to be in order to drive behavior quasi-inferentially. That’s one lesson of the finding that, under load, Joe’s behavior is not discordant with his explicit belief. I suggest that the thought causes his behavior deviantly. It generates anxiety, and the anxiety, in turn, non-deviantly causes the behavior discordant with his explicit belief.

People sometimes describe NA cases as mild or non-clinical forms of OCD. The comparison between these cases and true OCD is instructive, and may enable us to better understand how representational content is poised in the former.

There is some evidence that overavailability of motor representations to consciousness plays a role in the etiology of token episodes of OCD. For instance, there is evidence that information pertaining to habitual actions that is processed in nonconscious cortico-striatal networks in healthy controls is instead processed in fronto-hippocampal circuitry normally involved in conscious processing in patient groups (Milad and Rauch 2012). This overavailability may produce abnormalities in the phenomenology of habitual actions for OCD sufferers. This overavailability, in turn, prompts heightened attention to the routine behavior. Perhaps the causal route is best explained by some kind of predictive coding account, with the initial anomalies constituting surprisal and the allocation of attention to the anomaly an automatic response which has the function of making the prediction error more precise (Hohwy 2012; Clark 2013). The phenomenological anomalies may lead the sufferer to a sense of estrangement from their own habitual actions, and this may in turn lead to a decrease in their trust of the habitual actions’ efficacy (on an account of this sort, the rituals in which sufferers engage would be aimed at minimizing the prediction error).

Phenomenological anomalies and allocation of attention play no such role in NA cases. Whereas conscious attention to one’s actions does not alleviate OCD symptoms, and may actually be involved in their generation, conscious attention to one’s actions does not generate NA cases, and may instead lead to the cessation of anxiety. Nothing is amiss with agents like Joe until he leaves the house. His routine actions are phenomenologically normal; were there anything amiss, he would have attended to them at the time, rather than only after he left. Further, had he paid special attention to his routines as he acted, either prompted by some phenomenological anomaly or for some other reason, so doing would have led to the formation of a distinct dated memory of performing the action (say, turning the stove off), and that would have preempted the formation of an NA case. Whereas the sense of estrangement from routine behavior leads OCD sufferers to a reduced level of trust in its efficacy, NA cases feature nothing that would explain such estrangement.

Joe’s problems begin only when (for some reason) he finds himself imagining that he has left the stove on.\(^6\) Joe doesn’t mistake his imag-
Have I Turned the Stove Off? Explaining Everyday Anxiety

Neil Levy

in a belief, in the manner Currie (2000) describes, nor is his imagining unscrutinized by DLPFC mechanisms that test thoughts for consistency and intersubjective acceptability, à la Gerrans (2014). Correlatively, perhaps, Joe’s imagining does not play the belief-mimicking role Currie and Gerrans attribute to imaginings and thoughts. Rather, it impacts on his behavior by a deviant causal route. For some reason, imagining that he has left the stove on generates anxiety in Joe today, and that anxiety leads to rumination and possibly other belief-discordant behavior. 7

Joe’s anxious thought has representational content. He imagines that he has left the stove on (or something in that ballpark). But his thought is not poised to cause behavior in a belief-like manner. Were it poised to cause behavior in that way, there would be releasing conditions which lead to behavior. That’s precisely what we see with implicit attitudes, aliefs and other states akin to them. Under load, agents find it difficult to inhibit action tendencies associated with such states. Joe may perform actions that match the kind of content we attribute to his anxious thought (ruminating whether the stove is on, or even turning back and checking it), but because he is not more strongly disposed toward these behaviors under load, we have reason to doubt that the thought causes these matching actions non-deviantly. Instead, it cause anxiety, and the anxiety (non-deviantly) causes behavior.

Though there is an apparent match between the content of Joe’s thought and the behavior in which he engages, the match is only approximate. If Joe, unable to dismiss his worries from his mind any other way, returns to his apartment, he does not do so to check whether the stove is off; I suggest, but rather to assuage his anxiety. The difference in his goals will be evident in his dispositions. Someone who returns to his apartment to check whether the stove is off (like the person with genuine OCD) will be disposed to feel much less surprise if it is in fact on than someone who returns to assuage his anxiety. Joe, who assigns very little credence to the proposition that the stove is on, is disposed to be shocked if he finds it on. That’s an indication that his goal is not checking on the stove, and that, in turn, is an indication that his thought that the stove is on is not poised to cause behavior non-deviantly.

Other accounts of discordancy cases often explain them by reference to exotic states and processes (aliefs, say)—that is to say, states and processes unrecognized by folk psychology. The account preferred here postulates nothing very exotic. Neither the causal process it postulates, non-deviant though it is, nor the states in question should strike the folk as highly counterintuitive. It is no part of my agenda here to suggest that folk psychology can explain the full range of discordancy cases (indeed, elsewhere I have offered a somewhat exotic account of some of these cases myself [Levy 2015]). However, it may be that non-deviant causation by the kinds of states recognized by folk psychology explains some discordancy cases beyond NA cases.

Skywalker-style cases, for instance, might sometimes be explained by agents imagining that the Skywalk is unsafe (rather than alieving, or half believing, that it is unsafe). Imagining that might be unnerving, even in the absence of any temptation to believe what is imagined, and might therefore generate sufficient anxiety to (deviantly) cause behavior aimed at alleviating it: say, retreating to solid ground. Anxious thoughts might also play a (rather more convoluted) role in some kinds of departures from explicit egalitarianism, or in cases like Schwitzgebel’s trembling Stoic. While I doubt that anxious thoughts can influence Implicit Association Test scores in the envisaged manner, anxious thoughts might cause behavior that looks, and maybe is, subtly racist. For instance, the white person who imagines herself as possibly racist might be sufficiently worried by that thought to

7. On another day, Joe might have failed to entertain the thought that he had left the stove on, and even if he had, the thought usually would not cause him to feel anxious enough to generate belief-discordant behavior. Why can’t he dismiss the thought today? While I lack a satisfactory answer to this question, my own experience suggests that NA cases are more likely to arise when there is background anxiety concerning, or a heightened sense of responsibility for ensuring, the safety and security of something valuable. Departing for a long trip, and thereby leaving my apartment vulnerable, or having a friend’s house temporarily in my care are situations in which NA cases seem more likely to arise for me; these are situations in which I am disposed to feel more anxiety, so it is unsurprising that a passing fancy that I would easily dismiss under other circumstances is harder to dislodge in them.
become stilted in her behavior in the company of black people. Similarly, the wavering Stoic might waver not because he actually is not fully convinced by Stoic arguments, but because he represents himself as unconvinced and is made anxious by that thought. I doubt that the differences between the different kinds of cases (alief, in-between belief, imaginings taken for belief, and anxious thought) are always—or even often—introspectible. Neither we, nor observers, may be able to tell whether we are in-between believers or anxious thinkers in particular cases. But the differences between cases are nevertheless genuine and, in principle, can be made manifest: under cognitive load, for instance, the anxious thinker’s behavior will better reflect her explicit attitude, while load manipulations will have the opposite effect on those who alieve or in-between believe what they explicitly deny.

Conclusion

NA cases are discordancy cases, because agents like Joe behave in ways that conflict with their professed beliefs. They may engage in overt behavior, like returning to check the stove; at very least they engage in mental actions like rumination which conflict with their beliefs. But NA cases are very different from other discordancy cases. Whereas other cases seem to involve representational states with conflicting contents (aliefs or conflicting beliefs or partial beliefs), which occupy functional roles similar to those of beliefs, NA cases involve representations that do not occupy such roles. While self-deception and other discordancy cases might be motivated by affective states, NA cases alone are proximately explained by them: it is the accompanying affect rather than the representational content that non-deviantly causes behavior in these cases. The representational content of the thoughts entertained by agents like Joe deviantly causes their behavior, by generating anxiety that motivates belief-discordant behavior.

Explaining NA cases does not require the postulation of any mental states or processes other than those countenanced by folk psychology. I do not doubt that folk psychology is unable to account for the full range of discordancy cases, but, given that we ought to avoid multiplying mental states unnecessarily, we need to ensure that we are postulating exotic states and processes only when they are truly needed. The resources provided by the account offered here may explain some discordancy cases other than true NA cases, and thereby better delimit the set of cases that require more exotic explanation. Beyond that, I think the account has an interest in its own right, insofar as it helps us to understand some of the kinds of irrationality to which bounded beings like us are subject.²

References


8. I am grateful to an audience at the University of Birmingham and to two anonymous referees for this journal for helpful comments on an earlier version of this paper.


