

## Some Questions for Tamar Szabó Gendler\*

Forthcoming in *Analysis*

Tyler Doggett

University of Vermont

tdoggett@uvm.edu

*Intuition, Imagination, and Philosophical Methodology* is a superb book containing some wonderful essays. What follows asks questions mainly about three.

### 1. Self-Deception as Pretense

In the essay of the same name, Gendler defends a thesis she calls “SELF-DECEPTION AS PRETENSE” according to which,

[T]he cleanest and most interesting cases of self-deception share the following characteristic: ... A person who is self-deceived about not-P pretends that not-P is the case, often while believing that P is the case and not believing that not-P is the case. The pretense that not-P largely plays the role normally played by belief in terms of (i) introspective vivacity and (ii) motivation of action in a wide range of circumstances. (158)<sup>1</sup>

For illustration, consider Alan, self-deceived about the quality of his work. He tells anyone who will listen that it’s great stuff, though it should be apparent to anyone who reads it that it’s trash. He badgers OUP to put out a volume of his collected papers, though it’s apparent that OUP doesn’t publish the collected papers of no-hopers. Your attempts to get Alan to admit his work is no good are shrugged off as signs of your small-mindedness or jealousy. Nevertheless, Alan is a man of some taste and insight: Evaluating others’ work, he makes good—and uncontroversial—judgments of better and worse. He knows what OUP’s standards are.

I assume this is among the cleanest and most interesting cases of self-deception. If so, then, on Gendler’s view, Alan pretends his work is great. That is, he imagines that his work is great and this imagining—somehow; the details are unimportant—motivates him to tell everyone it is great, badger OUP, etc. Furthermore, the imagining makes the proposition that his work is great quite vivid to him. So his work seems great to him, and he acts like it is great, but this is not because he believes it is great. Rather, it is because his imagining that his work is great does what the belief with that content typically does. On Gendler’s view, Alan might well believe his work is trash and not believe otherwise.

Is there more to self-deception? Gendler does not claim that the characteristic highlighted by SELF-DECEPTION AS PRETENSE suffices for self-deception. (The qualification about “cleanest and most interesting cases” indicates that it isn’t necessary for self-deception.)

Whatever its status, the characteristic highlighted by SELF-DECEPTION AS PRETENSE is supposed to get at something especially important about self-deception. (Certainly the name Gendler gives the thesis suggests this.) It strikes me as an extremely intriguing and promising view. Its virtues are many. Questions in the rest of this section ask about just

---

\* I thank Andy Egan, Elizabeth Harman, and Arthur Kuflik for help with what follows.

<sup>1</sup> A closely related view is discussed in Darwall (1988) and attributed therein to Kent Bach and, perhaps, Bishop Butler.

Unless otherwise noted, all page numbers are from Gendler (2010).

one Gendler claims for it: She holds that SELF-DECEPTION AS PRETENSE “accords well with the sense that self-deception is an irrational condition” (159).

To test this, consider Bob, aged six, obsessed with trolls. He acts like a troll—well, like how he is told they act—talks like a troll—well...—etc. To his family’s frustration, he does not stop acting like a troll. When he thinks about what he is, he stays within the pretense and thinks that he is a fearsome troll, that he is not a person, etc. When he thinks about what to eat, he thinks of rotting flesh, etc. The pretense that he is a troll—and not a person—plays the role normally played by belief in making the proposition that he is a troll vivid to him and in motivating action. And it motivates a wide range of actions since all day he’s a troll.

Whatever his failings, Bob isn’t irrational. Annoying. Obstinate. Childish. Not irrational. In his case, there is nothing irrational about having his imagining be overwhelmingly vivid or letting his behavior be guided by such an imagining. He is immersed in an imaginative project.<sup>2</sup> There needn’t be anything rationally off about such immersion. Bob, like Alan, “pretends that [P] is the case, ...while believing that [not-P] is the case and not believing that [P] is the case” and his “pretense that [P] largely plays the role normally played by belief in terms of (i) introspective vivacity and (ii) motivation of action in a wide range of circumstances”<sup>3</sup> (158), but because Bob, unlike Alan, is not irrational, I worry that SELF-DECEPTION AS PRETENSE leaves out what’s distinctively irrational about self-deception and so I worry that Gendler is wrong when she says SELF-DECEPTION AS PRETENSE “accords well with the sense that self-deception is an irrational condition” (159). At least, I believe she is wrong *if* she means not only that the view is consistent with an explanation of the irrationality of self-deception but, to boot, provides that explanation.

Perhaps she doesn’t mean this. Gendler writes

[My] account accords well with the sense that self-deception is an irrational condition: after all, one canonical characterization of irrationality is that it is a state where something imaginary *inappropriately* comes to play the cognitive role of something real. (159; emphasis added)

That appropriateness condition is missing from SELF-DECEPTION AS PRETENSE. We could build it in so that the thesis says, roughly, that in the self-deceived, the pretense that not-P largely *and inappropriately* plays the role normally played by belief in terms of (i) introspective vivacity and (ii) motivation of action in a wide range of circumstances. Doing so would raise the questions of what it is for the imagination to *inappropriately* play a belief-like role and when it does so. Why, for example, is the role imagination plays in Alan’s actions inappropriate but the role it plays in Bob’s actions not inappropriate?

Gendler hints at one inappropriate role when she writes,

If the self-deceived subject were to reflect on her cognitive state (without inquiring too deeply), it would likely seem to her that she *does not* believe that P (after all, its normal manifestations are absent), but that she *does* believe that not-P (after all, it is playing a typically belief-like role). That is, the self-deceived subject may well have a false belief about what she believes: she believes that she believes not-P, while in fact she does not... In this sense, self-deception involves deception—or ignorance—about the self; it rests on

---

<sup>2</sup> On immersion, see Liao (2011), Schellenberg (2011), Velleman (2000a), and the references therein.

<sup>3</sup> So that condition does not suffice for self-deception since Bob is not self-deceived.

a certain failure of self-knowledge, or at least self-awareness. (170, fn. 38; cf. Holton (2001))

So self-deceived Alan's imagining that he is a talent makes the proposition that he is a talent quite vivid and motivates him to act like he is a talent. As a result, he becomes confused about what he believes. He falsely believes that he believes he is a talent. By contrast, when Bob reflects on his cognitive state, even thoughtlessly, he knows he believes he is a person and knows he doesn't believe he's a troll. While the content *You are a troll* is quite vivid to him, he is not tempted by the idea that he believes it. So perhaps a key difference between them involves Alan, but not Bob, being screened off from knowing about his own beliefs, at least insofar as he doesn't inquire too deeply. The pretense does the screening in Alan's case but not in Bob's.

Perhaps. Is the inappropriate role imagination plays in Alan's mental life inappropriate that imagination hides his beliefs from him? That it makes various propositions that he does not believe quite vivid to him? And is it crucial to the cleanest and most interesting cases of self-deception that the agent has false beliefs about beliefs?<sup>4</sup>

## 2. Self-Deception as Alief

Although all three were published in 2008, "Self-Deception as a Pretense" came into being several years before the two papers in *Intuition, Imagination, and Philosophical Methodology* on the mental state Gendler dubs "alief". In those two papers—"Alief and Belief" and "Alief in Action (and Reaction)"—Gendler explains that alief is a mental state comprising three bits:

1. A representation that something is the case
2. Some affect
3. Some behavior preparation

Furthermore, the three bits are "co-activated" (263 and *passim*) as follows: Typically, agents who go in for the representation also go in for the affect and behavior preparation, and the affect and behavior preparation are responses to the representation (263-264, 288-289). For example, seeing a snake nearby produces a certain representation that a snake is nearby and, in turn, that typically produces a feeling of discomfort and a move to flee.

Some representation-affect-behavior preparation triads—such as the snake alief just mentioned—are the same across a wide range of persons. Other triads are person-specific.

---

<sup>4</sup> In her introduction to *Intuition, Imagination, and Philosophical Methodology*, Gendler suggests that origin of Alan's imagining that he is a talent is also important:

[In a case in which someone is self-deceived about not-P], the self-deceived subject's attitude towards not-P is, in important ways, reality-indifferent: she holds not-P not because it is (taken to be) true, but because she wishes to be...in a not-P world. (11; cf. 171))

This sounds plausible. It, like SELF-DECEPTION AS PRETENSE, makes good sense of that fact that it is common enough for cases of self-deception to start life as cases of other-deception: Alan acts like his work is great in order to get others to come around to that belief. Chris acts like his partner is faithful because he doesn't want his neighbors to think otherwise. Etc. The pretense then somehow—to use one of Gendler's nice images from "Imaginative Contagion" (238-254)—infects other mental states with the result that Alan and Chris wind up self-deceived.

The wishful thinking needn't distinguish Bob from Alan. Bob might imagine he is not a person and that he is a troll because he wishes to be a troll and not a person. There is no problem for Gendler here—she clearly does not put forward the point about origin as a sufficient condition for self-deception—just a request for more information about the etiology of self-deception.

If, say, water bottles terrify me, it might be that a certain representation that there is a water bottle in front of me produces a feeling of discomfort and a move to flee. But, of course, that representation would not in most persons produce discomfort and motivate flight.

Finally, alief is habitually induced by certain stimuli, though which stimuli do the triggering varies somewhat from person to person.<sup>5</sup>

Illustrating alief, Gendler describes a situation in which she leaves her wallet at home and begs some money from a friend. After getting the cash, she absent-mindedly reaches into her bag as if to put the cash into her wallet. Describing what goes on here, Gendler claims,

The visual-motor input associated with handling cash rendered occurrent my...alief with the content: ‘Bunch of money. [This is the representation] Needs to go into a safe place. [This is the affect, I think.] Activate wallet-retrieval motor routine now. [This is the behavior preparation]’—an alief that ran counter to my explicit belief that my wallet was in Connecticut while I was in Maryland. (262; though Gendler notes that her use of “content” perhaps stretches the bounds of the notion of content (256, fn 3), I follow her terminology)

That is, seeing the money put Gendler into a state in which she represented that she had a lot of cash, felt the need to safe-guard, and readied a safe-guarding routine. That state is alief.

Further illustrating alief, Gendler recounts the results of a 1990 study by Paul Rozin in which subjects placed what they knew to be sugar in two containers. On one, they affixed the label “sucrose, table sugar.” On the other, they affixed the label, “not sodium cyanide, not poison.” Subjects showed considerable reluctance to eat from the latter. What went on in such subjects? Gendler speculates that, despite the not-cyanide prompt, they were moved by an alief with the content “cyanide [that’s the representation], dangerous [that’s the affect], avoid [that’s the behavior preparation]” (268-269). Besides helping to exemplify alief, the example reveals something about alief’s representational bit: Seeing the not-cyanide prompt, subjects nevertheless represented that there was some cyanide before them.

Gendler uses alief to explain various puzzling—and various unpuzzling—features of our mental lives and actions. She says that she would explain many examples of behavior purportedly motivated by familiar folk-psychological states—belief, desire, imagination—in terms of alief, including some examples from her earlier work. Should self-deception be so explained?<sup>6</sup> If so, then instead of SELF-DECEPTION AS PRETENSE, Gendler might endorse

SELF-DECEPTION AS ALIEF—The cleanest and most interesting cases of self-deception share the following characteristic: A person who is self-deceived about not-P represents in the distinctly alief-like way that not-P is the case, often while believing P is the case and not believing that not-P is the case. The alief including the representation that not-P largely plays the role normally played by belief in terms of (i) introspective vivacity and (ii) motivation of action in a wide range of circumstances.

---

<sup>5</sup> In the first alief paper, Gendler defines alief as a mental state. In the second paper, she says it is “to a reasonable approximation, an innate or habitual propensity to respond to an apparent stimulus in a particular way” (282, 288). The second characterization is repeated in the introduction to Gendler (2010) and in Gendler (2011a) though her view remains, in the second paper and in the introduction, at least, that aliefs are mental states (14, 288).

What are the (interesting) relations between the state and propensity accounts?  
<sup>6</sup> Gendler rejects explaining alief in terms of self-deception (259-260).

And if so, then what is goes on with Alan, self-deceived author, is that he represents that his work is great and acts like it is. However, the representation isn't belief—it's the representational component of alief—and the action is motivated by the behavioral preparation done by alief. Alan might well *believe* his work is no good, though his alief might hide that belief from view: "it would likely seem to [him] that [he] *does not* believe that [his work is no good] (after all, [that belief's] normal manifestations are absent), but that [he] *does* believe that [his work is great] (after all, [his alief that it is] is playing a typically belief-like role)" (170, fn 38; I changed 'pretense' for 'alief'). What goes on with Bob, young troll, by contrast, is simply pretense. Bob does not alieve that he is a troll since—and I will return to this—the representational bit of alief is not imagination.<sup>7</sup> So the characteristic SELF-DECEPTION AS ALIEF highlights—unlike the characteristic SELF-DECEPTION AS PRETENSE does—is had by Alan but not by Bob.

Importantly, SELF-DECEPTION AS ALIEF seems to me better placed than the SELF-DECEPTION AS PRETENSE to handle the worry about rationality I raised in §1. Gendler writes,

If I believe that P and imagine that not-P, I am violating no norms. But if I believe that P and alieve that not-P, something is amiss. Learning that not-P may well not cause me to cease alieving that P—but if it doesn't, then I'm violating certain norms of cognitive-behavioral coherence. No such criticism is possible in the analogous case of imagining. (271; cf. 14-15)

That is, Alan is not susceptible to criticism on grounds that he believes his work is no good and imagines it is great. Since there is nothing amiss with the imagining, I assume there is nothing amiss with Alan's pretense. It is because no such criticism of pretense is possible that SELF-DECEPTION AS PRETENSE is vulnerable to the criticism I raised in §1. By contrast, since Alan *is* susceptible to criticism on grounds that he believes his work is no good and alieves that it is great, SELF-DECEPTION AS ALIEF might do better on this front. However, this point about the criticizability of alief-motivated behavior raises two questions of its own: What are the norms of cognitive-behavioral coherence Gendler mentions above? Why does alief-motivated behavior differ from pretense in violating these norms?

### 3. Alief

Regardless of whether she explains self-deception in terms of alief, Gendler offers alief-based explanations of a variety of cases. In what follows, I focus on the wallet case already introduced, but the two alief essays feature many more.<sup>8</sup>

All agree that, in the wallet case, you represent to yourself that you have a lot of cash and, as a result, feel the need to safe-guard it and, as a result, start a safe-guarding routine. So all agree that, in the wallet case, there is a representational state and that affect and behavioral preparation are produced by it. But, presumably, not all agree there is alief in that case. One way to spell out a disagreement with Gendler: She is wrong about the identity of alief's representational state. She claims that that representational state is neither imagination nor belief, and perhaps she is wrong about that. Gendler makes three main arguments for

---

<sup>7</sup> One way to see Bob is no aliever: One can't alieve at will (271), but Bob pretends he is a troll at will. One can, in general, imagine at will. But only in general. On the limits of imagining at will, see Gendler's "The Puzzle of Imaginative Resistance" and "Imaginative Resistance Revisited" (179-226).

<sup>8</sup> See pp. 256-258 and 286 for some lists.

the representational bit of alief not being belief or imagination: One stems from a premise about opacity of contexts (270); another stems from Rozin’s sugar case (268-270); and the last stems from the primitiveness<sup>9</sup> of the representational state compared to imagination and belief (288-289, 288 fn. 8). As the page numbers just cited suggest, each argument is quite brief, and some will balk at their conclusion. But even those convinced by the arguments about what the representational state *isn’t* might be left wondering what it is: The nature of the representational bit of alief is somewhat underdescribed. It isn’t belief or imagination. It “operate[s] without the intervention of conscious thought” (288) and represents “some object or concept or situation or circumstance, perhaps propositionally, perhaps non-propositionally, perhaps conceptually, perhaps non-conceptually” (263-264). What else?

Even if Gendler is right about the identity of the representational state and about co-activation, her view doesn’t yet follow. Her view is not simply that there is a representation and it produces some affect and some behavior preparation. Rather, it is that there is a state comprising these three bits, and, furthermore, that state is not simply a hodgepodge of the three. To use some imagery from Leibniz, alief is sheep rather than flock.<sup>10</sup> To use some imagery from David Lewis, it’s trout rather than trout-turkey.<sup>11</sup> This leads to another way to spell out a disagreement with Gendler about alief:

...[A] potential worry [about alief]:...alief is not a fundamental mental state, but instead an *amalgam* of several more primitive mental states: those of entertaining content R, experiencing affect A, and activating behavioral repertoire B. I reply: the fact that our current vocabulary requires us to describe alief-content using three separate terms doesn’t show that the state is an amalgam of three others. (264, fn 16; emphasis added)

The worry Gendler responds to here can be developed like this: Everyone agrees that in the wallet case, there is a representation of something and, as a result, some affect and behavior preparation. Assume we all come to agree with Gendler about the identity of the representational bit. If so, the objection goes, there is nothing else to alief: There are these three bits, two produced by the third. And I take it Gendler’s response—in the second sentence, at least, of the passage above—is that there is something more to it, namely, whether there is some state comprising the three *and* where this state is *not* simply an amalgam of the three. Gendler rightly points out that the worry might simply arise because we currently describe alief so it appears to comprise three discrete bits. Nevertheless, one might wonder if the worry will persist even as our descriptive powers improve. Is there a unity to alief where this goes beyond the fact that typically there is the co-activation of alief’s components? If so, is the unity doing explanatory work for Gendler? What?

#### 4. Belief

The passage quoted above continues importantly:

The fact that our current vocabulary requires us to describe alief-content using three separate terms doesn’t show that the state is an amalgam of three others. *Indeed, one might even think that it is out of these more primitive association*

---

<sup>9</sup> Gendler argues that the representational state is more primitive than at least belief and imagination in at least two ways: It is present in cognitively unsophisticated creatures (283, 285, 286, 288-289, 290); and it is present in our evolution before belief and, presumably, imagination (288 and 288, fn. 8). I discuss a third way in §4.

<sup>10</sup> Leibniz (1989): 79.

<sup>11</sup> Lewis (1991): 80.

*patterns [the bits comprising alief]...that the less fundamental differentiated attitudes like belief, desire, and imagination are constructed. (Ibid.; emphasis added)*

The italicized text includes the hypothesis that belief, desire, and imagination are built from the components of alief. This hypothesis would explain two curious features of the alief papers. First, on one hand, Gendler writes, “I...argue that any theory that helps itself to notions like belief, desire, and pretense needs to include a notion like alief in order to make proper sense of a wide range of otherwise perplexing phenomena” (262) and writes, “I am making a parity argument:...any theory that makes appeal to...belief, desire, and pretense in order to explain behavior needs to make appeal to (something like) alief” (285). On the other hand, her arguments about alief typically conclude that there is alief and, importantly, her explanations of action typically don’t invoke alief *plus* belief, desire, imagination, etc. Rather, in these explanations, alief does all the explanatory work.<sup>12</sup> So why the parity claim? The hypothesis provides an answer: Any theory that includes belief, desire, and imagination needs to include alief because these other, more familiar states are—must be?—made of the bits of alief.

Second, Gendler characterizes alief as “conceptually antecedent” to belief, desire, or imagination or any other “cognitive attitude that [a] creature may go on to develop” (288). When she explains what she has in mind by “conceptually antecedent,” she writes, “Aliefs are more primitive than beliefs or desires. While it may be possible to paraphrase the content of alief using the language of belief and desire, alief cannot be factorized into belief and desire” (*ibid.*). If Gendler were arguing that aliefs are more primitive because not composed of beliefs or desires, her argument would be obviously fallacious. But that needn’t be the argument. Maybe Gendler’s view is that alief is not belief plus this or that or desire plus or imagination plus or...; by contrast, as the hypothesis implies, these more familiar states are alief plus this or that. It is in that way that alief is antecedent to the more familiar states.

So, finally, Gendler says *one might* think that belief, desire, etc are constructed, at least in part, out of the bits of alief. *Should we* think that?

## 5. Conclusion

As my profusion of questions probably makes clear, I found *Intuition, Imagination, and Philosophical Methodology* extremely thought-provoking. It is one of the most stimulating works of philosophy—or anything else—I have read in years.

## Works Cited

- Darwall, Stephen. 1988. “Self-Deception, Autonomy, and Moral Constitution” in McLaughlin and Rorty (1988).
- Doggett, Tyler. 2011. Review of Gendler (2010). *Notre Dame Philosophical Reviews*.
- Egan, Andy. 2011. “Comments on Tamar Gendler’s ‘On the Epistemic Costs of Implicit Bias,’” *Philosophical Studies* 156: 65-79.
- Gendler, Tamar Szabó. 2010. *Intuition, Imagination, and Philosophical Methodology*. New York: Oxford University Press.
- Gendler, Tamar Szabó. 2011a. “On the Epistemic Costs of Implicit Bias,” *Philosophical Studies* 156: 33-63.

---

<sup>12</sup> I go in to some detail about the arguments for alief in Doggett (2011).

For more on the explanatory usefulness of alief see Egan (2011), Gendler (2011a), Gendler (2011b), Haug (2011), Kriegel (2011), and Kwong (2011).

- Gendler, Tamar Szabó. 2011b. "Moral Psychology for People with Brains," manuscript.
- Haug, Matthew. 2011. "Explaining the Placebo Effect: Aliefs, Beliefs, and Conditioning," *Philosophical Psychology* 24:679-698.
- Holton, Richard. 2001. "What is the Role of the Self in Self-Deception?," *Proceedings of the Aristotelian Society* 101: 53-69
- Kriegel, Uriah. 2011. "Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction," *Australasian Journal of Philosophy* (forthcoming).
- Kwong, Jack M.C. 2011. "Resisting Aliefs: Gendler on Alief-Discordant Behaviors," *Philosophical Psychology*, forthcoming.
- Leibniz, G.W. 1989. *Philosophical Essays*. Indianapolis: Hackett.
- Lewis, David. 1991. *Parts of Classes*. Oxford: Blackwell.
- Liao, Shen-yi. 2011. "Becoming Immersed," manuscript.
- McLaughlin, Brian and Rorty, Amelie, ed.s. 1988. *Perspectives on Self-Deception*. Berkeley: University of California Press.
- Schellenberg, Susanna. 2011. "Belief and Desire in Imagination and Immersion," manuscript.
- Velleman, J. David. 2000a. "The Aim of Belief" in Velleman (2000b).
- Velleman, J. David. 2000b. *The Possibility of Practical Reason*. New York: Oxford University Press.