

Keeping Self-Deception in Perspective

Lawrence Beyer

If one accepts the suggestion that self-deception involves simultaneously held beliefs that are contradictory (that *p*, and that not-*p*) or at least inconsistent (that the evidence favors '*p*,' and that not-*p*), then some special structure must be postulated that allows these to coexist in the same mind without either extinguishing the other. Donald Davidson sees this function served by a partitioning that bounds different, though overlapping, territories within the mind and keeps their contents from coming into contact (e.g., 1998, 8).

Mark Ansbach, however, expresses Sartre-inspired misgivings about this picture. Taking psychotics (e.g., "the sort of individuals who believe they are... Napoleon" (1998, 60)) to exemplify "extreme self-deception" (61), he claims that self-deception "involves" and "requires" "the simultaneous presence of contradictory beliefs" (60),¹ produced by a *paradoxical* kind of resistance to a proposition known, believed, or suspected to be true. But Ansbach says little about the central beliefs constituting the self-

1. In general, it is not clear whether Ansbach's references to 'self-deception' and 'bad faith' ought to be taken to refer only to the contradiction form of self-deception, or whether his account can be generalized to cases involving opposed-but-not-contradictory beliefs (i.e., that the evidence argues that *p*, and that not-*p*).

deception (e.g., 'I am not Napoleon,' 'I am Napoleon'); his focus is one level removed, upon self-awareness of that condition. Such awareness is available to all self-deceivers, he suggests (73, 80–81), but only within an "oscillatory structure" shared with unawareness (77, 81). The conflicting second-order beliefs (or belief plus nonbelief) coexist together stably, not as if partitioned apart, but as if united though distinct on the double-yet-single side of a Möbius strip (79, 82–83). (Regarding madness specifically, the unresolvable oscillation is said by Anspach to block the self-awareness path to a cure.)

88 My agenda for this paper is an amalgamation of several lines of inquiry. After examining, in general terms, the gratuitousness of self-deception attributions, I cast doubts upon Davidson's arguments that lying to oneself is impossible and so cannot be what self-deception involves. This is followed by a rebuttal of Anspach's dismissal of partitioning accounts, and a critical analysis of his speculations concerning the "paradoxical" "oscillatory logic" (1998, 76–77, 81) of the self-deceiver's (or madman's) cognitive predicament. Finally, I provide two short appendices, the first clarifying one aspect of the distinction between self-deception and wishful thinking, the second offering a detailed explanation of why belief in one's own self-deception need not be paradoxical.

1. Attributing Self-Deception

Self-deception is not merely the possession of logically inconsistent beliefs, for these can be acquired and maintained innocently. Nor is it simply the adoption of a proposition as one's belief out of a desire for that event or state of affairs, for that may only be wishful thinking (see Appendix A). Self-deception seems to require that aversion to one already-present belief (that *p*, or that the evidence favors '*p*') motivate action that produces adoption of a contrary belief (that not-*p*).²

2. For purposes of this essay, I stick rather close in some ways to Davidson's basic models of irrationality. On Al Mele's (1998) account, by contrast, self-deception can be initiated not only by an aversion to believing that *p*, but also by a desire to believe that not-*p*, and these cause self-deception by instigating biased processes of evidence gathering, interpretation, and evaluation—thereby producing a belief that not-*p* without any intentional actions having been taken to bring this about. Thus, for Mele, wishful thinking may be only "a form of self-deception" (Mele 1987, 135) or perhaps a twin brother, by contrast with the wider Davidsonian distinction adopted in this essay.

But even that is not enough. For it is not self-deception if resistance to the original belief merely motivates a more thorough gathering or reexamination of evidence, and this then results in a change of belief. Self-deception demands that one's actions be *designed* to produce a change in one's belief, and that the reasoning process by which this is achieved be materially corrupted by this motive; for if one employs a scrupulously fair reasoning process that happens to achieve the intended cognitive outcome, one cannot be said to be self-deceived, even if one had been prepared to resort to underhanded methods if necessary.

Now, a threshold concern raised by these rather demanding conditions is whether self-deception actually exists—whether (as suggested by folk psychology and taken for granted by philosophers of mind) people really engage in such a thing, or, in other words, whether such attributions are ever either indispensable or at least the most fruitful explanatory hypotheses available.

In this section, I hope only to say enough about a 'perspectivist' alternative to impose upon philosophers a burden of justifying the attribution of self-deception, of explaining why this concept of folk psychology ought not be pared away by Occam's Razor. At the very least, there seems a disturbing readiness to diagnose self-deception prematurely, before exhausting other plausible but less epistemically unsavory possibilities.

Consider Anspach's paper as an illustration. He seems to reason that holding beliefs that strike others as "manifestly deluded" (1998, 60) entails being 'deceived'; and then, that unless someone else (or one's own DNA) has implanted these beliefs in one, one must be *self-deceived*.

This glib analysis of the concept of self-deception loses touch with the specific mental phenomenon described above. It also employs a spurious dichotomy, between extrinsically "implanted" versus "internally generated" beliefs (61), which subtly conveys the false notion that others' queer beliefs are insular fantasies and never reasonable responses to evidence.

Worst of all, the implicit criterion for self-deception (and *madness* for that matter) is disturbingly easy to satisfy: having beliefs that seem to other people, in light of the evidence as they interpret and evaluate it, to be obviously wrong or ridiculous. This alone is supposedly enough to make one self-deceived, and not just grossly in error or merely committed to a different cognitive framework or perspective. But as Davidson rightly insists (e.g., 1985, 348), one may simply have formed an interpretation of the world that makes sense given one's perspective (which may include

strange beliefs and theories), rather than a view motivated by aversion to an already-possessed contrary belief.

It is not easy to keep in mind that deviant beliefs can, in this way, be explained by mere difference in perspective. In fact, even relatively widespread evidence of another's perspectival divergence is all too readily construable simply as chronic, or perhaps metastasizing, self-deception. Given the natural assumptions that there is a single True account of reality and that users of a language (dialect) all share the same concepts, people readily reason, when they encounter another's judgment that differs from their own confident one, that both cannot be right; but, surely, theirs is right, so the Other's must be wrong. And since it appears that the Other is not stupid and had access to all the pertinent information that led them to *their* view, the explanation must be some irrational kink in the Other's stream of thought.³

But this is quadruply misguided. (i) Other persons often use words with senses and usage criteria subtly different from one's own, which can lend their avowed beliefs a false appearance of absurdity relative to the other beliefs or evidence they seem to possess. (ii) Thinking isn't easy; it's far from trivial to bring to bear in each judgment all of one's relevant reasoning skills and memories. (These include one's standards of cogency and weightiness in reasoning, explanation, and argumentation.) Even those with admirably efficient habits of mind are beset by (unmotivated) blind spots and slip-ups, and sometimes forget or overlook relevant considerations. (iii) Those evaluating another person's beliefs are typically handicapped by an information disparity: he has evidence unavailable to

3. Typical self-deception attributions do, in this way, involve matters that seem perfectly obvious to the attributors, and this contributes to the ease of unwarranted attribution. On occasion, however, a judgment of self-deception is not a response to a particular patently absurd belief, but an adjustment of one part of a multifaceted understanding of someone's mind in order to improve the logical fit of *other* attributions to that person. In such cases (which include self-deception *self*-attributions), the reevaluated belief need not seem on its own to be misguided, let alone obviously so. (There are other possibilities as well. For example, one might hold an attributional principle 'Whenever people believe their mates are perfect for them, they suffer from self-deception' and apply it categorically, even where one has no other reason to think that the beliefs in question are deluded.)

The atypical cases appear to be parasitic upon the typical ones, though, in that if the latter were no longer accepted as meriting the label 'self-deception,' the use of the concept could no longer be sustained in the former (and not vice versa). Hence my focus upon the central cases.

the smug onlookers, and often lacks evidence they possess. (iv) Most important: given the variability of perspectives, a slice of reality can be given more than one rational interpretation.

In sum, if we do not see someone's construal as a reasonable employment of his mental resources, the source of the difficulty may be *our* mental limitations, not his. Interpretive "charity" (e.g., Davidson 1982, 302–3) does not preclude restraint or retreat in the face of bafflement. The deepest attainable understanding of another being is often an acceptance that one cannot understand him in depth, that his perspective cannot be specified in enough detail, and that bending the pieces to get them to fit is no solution to the attributional puzzle he poses.

The same issues arise for *self*-ascriptions of self-deception that respond to the now-unfathomable beliefs of one's own past self. And a parallel explanation obtains when one has a *current* belief that from another standpoint one considers indefensible.

That is, one thinks in differing ways in differing types of situations. A somewhat altered frame of mind gets elicited or composed whenever one perceives (usually not self-consciously) a change in one's 'pragmatic context,' hence in one's immediate mental task, purpose, or activity. The differing frames of mind, which we might call 'perspects,' do not all make available the same full set of mental resources and dispositions. Each affords access to only a subset of one's memory, and thus lacks some intentional states available in other perspects.⁴

Hence one person may in different situations (perhaps very close in time) innocently use inconsistent beliefs about the same subject-matter, each one rationally acquired and maintained in the context of its surrounding perspect. (The beliefs may stably coexist in a state of *compatible inconsistency* if each is useful within its contexts of usage, if these domains do not overlap, and if no applicable norm demands that the beliefs be reconciled.) While using *one* "structure... of interlocking beliefs, expectata-tions, assumptions, attitudes and desires" (Davidson 1982, 300), one may be reminded of a belief that is part of another, and from the current

4. Such a frame of mind is, roughly, a subset of one's 'perspective' or total mental endowment, and it functions to provide a particular point of view, guiding and constraining mental processes—thus the name 'perspect.' According to this conception, a perspective, like a hockey or football team, is at no time ever fully activated and available for use. It is only through the mediation of a single perspect at a time that a person engages the world, and in that sense perspects are psychologically and explanatorily prior to perspectives.

standpoint that belief may seem ludicrous. So here again, a perspectivist explanation—one overlooked due to assumptions about the mind's unity or integration that obscure its structural complexity—seems superior to an ascription of self-deception

To posit a mind structured in this partly nonintegrated way is to posit a sort of mental partitioning, though one different from that proposed by Davidson.⁵ However, Davidson's own account of self-deception may itself rely tacitly also upon some such organizational arrangement. A self-deceiver who promotes a belief by "pushing the negative evidence into the background" (1986, 90), say, is not merely turning attention away from certain of his beliefs (and other intentional states), for that would provide no assurance against their prompt return. He is placing them out of bounds, making them unintrusive upon reasoning for the duration of the self-deceptive project—hence the seeming aptness of a 'partitioning' idiom. Once the new belief is in place, they may well be readmitted without threat, given the remarkable tendency to reconstrue data so as to harmonize it with established beliefs—now including the induced one. (This prodigious reinterpreted capacity also casts doubt upon Davidson's insistence that the aversive original belief and its motivating power must survive the self-deceptive inducing of the contrary belief lest the latter be smothered in its crib by remembered or freshly encountered evidence (1986, 90).)

Now, a plausible self-deception hypothesis in a particular instance must involve a strange belief(-set) that is both isolated (for the more wide-ranging an attribution of irrationality, the greater the suspicion that the overall interpretation can be improved upon) and best construed as *motivated* to evade an uncomfortable realization. However, determining what beliefs are so aversive to someone as to bring about self-deception only expands the opportunity for interpretive arrogance and disregard of perspectival difference. The spottiness and ambiguity of evidence insures that some story can always be constructed *ex post* about what the agent is afraid of facing—his sadness or, contrariwise, his happiness; infidelity, or fidelity that he does not deserve.

5. John Heil (1994) mentions mental "compartmentalization" of this type, and seems to consider it a situation of "mental regions functionally circumscribed," but apparently resists conceiving it as involving "mental divisions" or "partitions." Bill Talbott (1995, 29,34), by contrast, states unequivocally that certain "innocent divisions" are required for "adequate explanation of non-self-deceptive phenomena."

At times, the search for independent evidence of motivation gets skipped altogether. Anspach, for instance, shows no interest in asking why someone would be so disturbed by suspecting he is not Napoleon that he would contrive to believe the opposite. The explanation would likely point to some alien-to-us reasoning—but then parsimony would dictate attributing the 'I am Napoleon' belief to that alien body of beliefs alone, without self-deception. Anspach, though, infers self-deception (hence that some suitable motivation *must* be in play) simply from steadfastness in an enigmatic belief in the face of efforts at rational persuasion (64):

If... a reasoned attack on his deluded beliefs actually provokes the patient to cling to them all the more strongly, one might surmise that the patient is resisting recognizing something as true out of a belief, or at least a nagging suspicion, that it *is* true. This resistance would then qualify as self-deception....

Who is not vulnerable to such 'logic'? Are not our convictions commonly fortified in the process of being defended against reasoned criticism? Where opposing webs of belief confront one another, neither seeming flawed by its own standards more than its opponent seems to be, why *should* reasoned argument win one of them over? There are, moreover, many innocent reasons even for *emotional* responses to a reasoned attack upon one's beliefs. (For a start, one's interlocutor may reasonably strike one as tedious, arrogant, insulting, inconsiderate, self-righteous, simple-minded, dogmatic, or a threat to disseminate misinformed, harmful views.)

Diagnoses of self-deception can carry serious consequences, starting with distrust and disrespect for the target's viewpoint (and ending, in some contexts, with autonomy-infringing manipulations and controls). Perhaps some of the same leeriness now aroused by suggestions of group 'false consciousness' ought to be directed toward attributions of individual self-deception.

A cloud of suspicion lurks over such ascriptions even where the charge is 'confirmed' (and perhaps volunteered) by the target himself. His ignorance and unreliability aside, self-deception (if it exists) involves a complex configuration of mental states, not an isolable, introspectively discernible one. One can affirm the theoretical hypothesis, but cannot have special first-person means to validate it; self-consciousness of one's self-

deception is not to be had.⁶ Though such affirmations are commonly construed as true insights dispelling the fogs of self-mystification (see, e.g., Talbott 1995, 40, 58), these construals presuppose the truth of the self-deception hypothesis and thus cannot provide support for it. 'Confessional' affirmations may simply be alternative, perhaps false, readings of the evidence (and possibly themselves influenced by, say, wishful thinking).⁷ Even where the hypothesis of self-deception produces searching self-scrutiny, and adjustments in outlook, it is not thereby verified, any more than is a charge of selfishness or racism that sparks reflection and change.

Suppose perspectivist psychological explanations obviate the need to attribute self-deception. They might still be thought to impute irrationality of a different sort, and thus to offer no great gain from the standpoint of interpretive charity. They ascribe to agents interpretations and evaluations that, in having been shaped to conform with already-existing beliefs, arguably have been illegitimately biased, even if this is the normal course that understanding takes. Psychologists have discussed a "primacy effect," whereby beliefs based upon earlier-encountered information cause later data to be interpreted to fit them, and hence are themselves "revised insufficiently in response to discrepancies in the later-presented information" (Nisbett and Ross 1980, 172). Now, even if such assimilation of new data to old were irrational, substituting a pervasive and ineliminable element of psychological explanation for the more dubious and dangerous self-deception concept could itself be an improvement. And *is* such assimilation

6. Where the overt, supposedly-self-deceptive belief is firmly held, it seems clear that introspective self-awareness of one's own self-deception is unattainable. Where weakly held (as shown, eg, in one's feelings toward it, or dispositions to use it), the weakness itself might be introspected, but there is no similar way to perceive that the belief was induced via self-deception and is not simply an ordinary belief about which one has serious reservations.

7. Compare Talbott's (1995) discussion of the Fox, who, after finding that he cannot reach the sweet-seeming grapes, reconsiders and decides that they are green and sour after all, and Mary, who, after her lover leaves her, reconsiders and decides that he had long been heartlessly taking advantage of her but that she had been too blind to see it. The former case is treated as a movement from rational to motivationally biased belief, the latter as a movement from self-deception to rational clarity. But it could be that none of the four beliefs is irrationally biased, or that it is the two that Talbott treats as rational that are in fact biased, perhaps by wishful thinking. Self-deception interpretations, even if offered by the putative victims themselves (e.g., Mary), are question-begging and do not provide independent evidence that self-deception exists.

of evidence really irrational? It does not violate the "total-evidence principle" (said by Davidson (1986, 91) to be a norm all rational agents must possess), which calls for choosing the hypothesis best supported by all the evidence. For a violation requires that the evidence be construed as favoring a competing hypothesis (1986, 82), and the putative self-deceiver seemingly does not do this, however much observers think he should.

Perhaps assimilating evidence to existing theory contravenes another norm of rationality. Now, there is surely no norm that requires that one *always* canvass for alternative hypotheses and interpretations of evidence. But Davidson (1985, 347-48) suggests one that demands considering alternative hypotheses upon finding oneself forced to "invent strange explanations" for observed data. Other principles belonging to the same family would demand considering alternative hypotheses upon finding one's judgments at odds with those of experts or the consensus of reasonable, intelligent people, or upon recognizing that one needs to make a high-stakes, irrevocable judgment. Still, if one either does not comply with such norms (which are only 'optional' for rational agents), or does entertain the alternative hypotheses yet reaffirms one's original view, allegations of irrationality seem to be left without basis.

So what, then, do we say of someone who consistently takes himself to be Napoleon?—That he has an abnormal perspective, and need not be thought self-deceived.⁸ The bizarreness of his beliefs may be due to isolation from the people around him, rather than due to a narrowly motivated process like self-deception. Most people rely upon others far more than he in forming their judgments; they handle the inherent difficulty of understanding the world by making thinking in effect a collective activity. Two heads *are* better than one. Cooperative communication lessens neglect of relevant factors, provokes frequent corrections—and limits how far different people's perspectives can drift beyond the limits of mutual comprehensibility. (Some deviant patterns of construal and reasoning

8. He might, for instance, use 'Napoleon' idiosyncratically, reflecting an eccentric metaphysical view. 'Napoleon' could, say, denote for him not a unique individual (as it does for us), but something closer to a type—to him, he is not numerically identical with the Napoleon we all know, but is *another* Napoleon with similar properties. If we allow philosophers and physicists to believe in possible worlds and branching universes, and heed anthropologists' and historians' accounts of other cultures and of the deviants and innovators within our own, we must concede that psychotics' languages and metaphysics may differ from ours.

may be underlain by neurochemical factors. But the interpretations may nevertheless be reasonable *given* the abnormal brain processes, and quite free of any motivated irrationality.)

2. The Intentions of Liars

Whatever self-deception is, Davidson contends that it is not lying to oneself (1998, 3; 1986, 88). But his main arguments do not appear to compel this conclusion.

96

Only two intentions are indispensable for any liar, he says (1986, 88). First: "to represent himself as believing what he does not." This seems needed because a liar achieves his aim by exploiting the victim's reliance, in his own (the victim's) belief-formation, upon what he takes the liar to believe. Typically, the liar is granted epistemic authority by the victim, so that the latter comes to believe that *p* because he (wrongly) takes the former to do so. (Atypically, the victim may think that the liar forms systematically and dependably false judgments in this area, and so when the liar appears to believe that *p*, the victim believes that not-*p*.)

Now, Davidson's construal of this first requirement seems to need some slight modification, as he apparently takes 'representing oneself as believing that *p*' to involve communicating some sign whose literal meaning is '*p*' (1998, 3). But when one represents oneself as believing something, one takes into account and makes allowances for the conventional nonliteral interpretations that one's audience is likely to place upon one's words. Surely one represents oneself as believing that *p* when one utters 'Not-*p*' in the knowledge that one's hearer will detect the obvious irony or facetiousness in one's tone and manner. Under such circumstances, it would be a feeble and ridicule-worthy defense to maintain that one didn't actually lie because one believed that not-*p* and, after all, one *said* that not-*p*.

The objection to Davidson's argument, however, does not turn upon the precise characterization of what it is to represent that one believes something. The more important feature of his statement of the first intention of the liar ("to represent himself as believing what he does not") is that it is formulated in abstract terms of which the liar need not be aware. The liar need not think of his own 'representing' or 'believing.' He may simply intend, for example, to say that *p* (which he does not believe), in order to get the victim to believe that *p*; he need not grasp that lying has anything to do with a victim's perception of and reliance upon the liar's

state of belief. A liar must *in fact* represent himself as believing what he does not, and must intend to do something that (perhaps unbeknownst to him) *can be described as such a representing*; but he need not intend 'to represent myself as believing something that I do not

The second intention that Davidson considers essential, "to keep this [first] intention... hidden from his hearer," seems to make lying to oneself impossible and hence an untenable interpretation of 'self-deception.' But while a successful lie does demand that any intention to (in effect) misrepresent one's belief not be found out by the victim,⁹ this need not call for intentional action on the liar's part. He may just, for example, offhandedly assert a knowing falsehood, without any concern that his misrepresentational intention (or simply the fact that he does not hold what he says to be true) might be discerned. (If it seems that a liar *must* be concerned with this, think of liars who do not think their lying is wrong, or do not care whether it is, or have unquestioning confidence in the ease with which they will con their victims.) If later the possibility of being found out comes to the liar's attention, he can *then* take steps to hide the original misrepresentational intention, forming an intention to do certain things or to improvise so as to keep it hidden.

From the fact that success in an activity requires that not *p*, and maybe that one not intend that *p*, it does not follow that one must intend that not-*p*. For example, though in order to use a tool one needs for it not to break, it does not follow that one uses it intending not to break it or have it break, unless and until there is a specific concern about that possibility. Again, one's bridge hand may need to stay hidden from opponents, but one need not intend to keep it hidden if one was trained to hold one's cards in a way that happens, as an incidental and nonfocal feature, to keep them hidden. Know-how obviates the need for explicit control over various aspects of an action.

And consider feinting, which like lying involves intentional action de-

9. By contrast, a liar's more general intention to deceive can be found out without this invariably undermining the lie. The victim may still miscalculate what the liar really believes and end up in the belief state sought for him by the liar. Indeed, a cunning liar might intentionally have his victim discover the intention to deceive. (For example, though I believe that not-*p*, I tell you that *p*; I then (without you knowing that I planned this) let you discover elsewhere that I intended to deceive you and that I expected or contrived for you to discover this; you then figure that I sought to induce you to believe that not-*p*, and that I really believe that *p*; so you accept that *p*.)

signed to produce a belief that one does not share. A basketball player's fake (e.g., a step toward the basket, when he really intends in a moment to jump out and receive a pass) must be communicated to the defender, while the misrepresentational intention must be hidden. Yet the player need not actively hide anything; his routinized action pattern keeps the crucial facts hidden, but it can be initiated and executed without this feature being intended (or noticed). And if a player *is* aware that discovery of his intention would be disastrous, he typically will not take that possibility seriously enough to form a second intention expressly to hide the first.

So, any general impossibility in lying to oneself apparently lies not in "doing something *with the intention that* that very intention should not be recognized by the intender" (Davidson 1986, 88; emphasis added), for the italicized intention need not exist.

Might it lie, then, in the impossibility of doing something with an intention (the misrepresentational one) that *is in fact* not recognized by the intender? Stated so generally, this hardly seems capable of rendering lying to oneself impossible in principle, for people seem often to perform actions without at the time being aware of their intentions or later recalling them. True, some aspect of an intention may need to be retained over time, lest the action go off the rails, transmute into something else, or simply be abandoned. But it can suffice to remember what one is doing, without also remembering why. All but the most narrow of one's aims can safely be forgotten (at least in many cases, and at least until a decision arises that requires consideration of higher-level aims), as can the background precipitating circumstances of the original intention to act.

Thus, if the original intention is to do something with the aim or hope of getting oneself to believe that not-*p*—treating 'not-*p*' as one's working premise in reasonings, for instance, or gathering a body of strong evidence to support it—it seems that one might pursue this activity effectively without retaining awareness of one's broader aims or of the precipitating circumstances, such as an initial aversion to the thought that *p* (or the thought that one ought rationally to believe that *p*). If so, one can retain as much of the intention as is needed to sustain one's actions, without also retaining a project-imperiling awareness of the misrepresentational aspect of the original intention.

That an intention to lie to oneself has distinct aspects that can be given different psychological treatments also supplies the reply to Davidson's new impossibility argument (1998, 3). The supposed impasse is that in ly-

ing to oneself, the liar must intend (or simply need¹⁰) both (i) that his intention—to misrepresent what he believes, by asserting that *p*—go unrecognized by the victim, lest the victim see through the deception, and also (ii) that the intention be recognized by the victim, insofar as any lie requires that the liar be taken as making an assertion that *p*.

If this is a fair synopsis of the new argument, it seemingly cannot be valid, for if valid it would show the impossibility of lying *to others*, too. The problem appears to reside in a failure to distinguish two different aspects of the intention—assertion and misrepresentation—one of which could be (intended to be) recognized, while the other could be (intended to be) not recognized. Reconsider how a liar's primary intention could simply be to say that *p* (which he does not believe), in order to get the victim to believe that *p*. The intention to say that *p* might be (intended to be) discovered, with the non-belief in '*p*' (or belief in its falsehood) (intended to be) not discovered. (Such a distinction was also noted above regarding feinting.)

The second Davidson argument, then, seems to fare no better than the first. Does the analysis of self-deception as lying break down elsewhere? If yes, perhaps it is where lying demands that the victim accept the proposition *because* he thinks the liar believes it; for where liar and victim are one, it is not clear how such (granting of and) reliance upon epistemic authority, and upon one's own apparent belief, can take place.

On the other hand, a scenario like the following seems possible. One intends, as suggested earlier, to accept 'not-*p*' as a working premise in one's reasonings and to find support for it. Later, once certain aspects of the original intention have been forgotten, one then accepts 'not-*p*' because one remembers oneself using and trusting it. Even a brief time separation might in this way allow one to bifurcate functionally into 'relier' and 'relied upon.'

The case against the possibility of lying to oneself thus seems yet unproven.

10. Davidson frames his argument in terms of states of affairs the liar "intends"—here, that his (the liar's) intention in making his lying assertion be both recognized and not recognized by the victim. But this focus upon intending seems superfluous, as it is the factual impossibility of the victim's recognition-plus-nonrecognition that is the crux. It also seems misleading. Davidson apparently infers (1998, 3) that a liar "must intend" that *p* from the fact that *p* happens to be a precondition for the successful execution of his action; but this does not follow. (When one asserts something, one need not, e.g., intend that one's hearer not have a deafness-inducing stroke as one commences).

3. Partitioning as an Explanation for Self-Deception

Anspach approvingly reiterates Sartre's flawed argument that where there is motivated resistance to acknowledging one of two opposed beliefs—as

Sartre (1956, 90) assumes a partitioning model that includes a *censor*, an "autonomous consciousness" (94) who polices the mental boundary, keeping hidden beliefs hidden and turning the conscious mind away from evidence in their favor. The censor itself would need be self-deceived, he argues, thus only displacing the problem of explaining self-deception, not resolving it. But Sartre never does establish that there need be any entity beyond the distinct mental subsets themselves. For instance, the unconscious mind itself might somehow monitor its conscious counterpart, instigating resistance when necessary. Neither the censor (which would not exist) nor unconscious (which would be unconflicted) would then itself be self-deceived.

Partitioning theories do need to explain the precise mechanics of how a hidden belief could cause an opposite manifest belief to be maintained. But, *contra* Anspach's extension of Sartre, such accounts need not "call for a censor to keep the agent from surveying the whole" (Anspach 1998, 67, n. 5); nothing Anspach or Sartre says shows that "one cannot limit the scope of consciousness by positing... a wall inside the mind without also positing... an independent consciousness to mind the wall" (Anspach 1998, 67). A 'censoring' effect might be produced by the structural organization of the mind, such as a system of transition rules or switches regulating the activations of the partitioned areas.

Even if one were for argument's sake to accept the censor's existence, and that it "must know what it is repressing" and "apprehend... [it] *as to be repressed*" (Sartre 1956, 93), Sartre's argument would not go through. For it would not follow that this mental module must have "an awareness of its activity." Sartre relies upon the premise that "[t]o know is to know that one knows," but an animal may know what it is stalking, and know it as something to be stalked, yet without self-consciousness at all. Sartre's argument seems to entail that we have self-conscious awareness of doing every activity we do—which seems false. (Absorbed in driving, one appre-

11. Sartre is actually discussing 'bad faith,' which he views as lying to oneself. Anspach sees lying to oneself as only (an ingredient in) one kind of self-deception, but treats Sartre's comments as sometimes applying to self-deception in general.

hends the road reflectors as objects to be skirted without thereby gaining an awareness of this activity.)

But assume that there *is* a censor, conscious of its own activity. Now, its self-awareness is supposedly "the consciousness (of) being conscious of the drive to be repressed, but precisely *in order not to be conscious of it*" (Sartre 1956, 94), with the crucial upshot that the censor itself is in bad faith. Yet however this climactic quotation is paraphrased—e.g., "The censor must have awareness of being conscious of the repressed matter, and yet it is also crucial [for the self-deception to succeed] that the censor not be conscious of that matter"—the second clause seems simply false. It makes no difference what the censor knows. The repressive activity is aimed at keeping the conscious mind (or rational ego)—not the censor—in the dark.

Anspach's borrowing from Sartre thus does not come close to establishing that partitioning cannot account for ordinary self-deception.¹² Sartre's interesting but convoluted account is rife with missteps, builds upon a dubious premise about "the total translucency of consciousness" (1956, 89),¹³ and is a shaky basis from which to draw conclusions about self-deception. (See also Pears (1986, 74) and the thorough discussion by Allen Wood (1988).)

4. The Supposed Paradoxes of Self-Deception

4.1 The Self-Aware Madman

Anspach uses the plight of a self-aware psychotic to argue that no escape from self-deception is offered by the flickering self-awareness of that con-

12. Anspach also objects (1998, 73) that a posited mental "dividing line cannot explain... [the fact of inconsistent beliefs] if its own existence is deduced from the fact to be explained." Rather than being "circular reasoning," however, this is simply normal 'abduction' or 'inference to the best explanation.'

13. Sartre's idealization of the mind imputes self-awareness where none exists (given his "dogma that *all* mental activity must be conscious" (Wood 1988, 211))—sometimes even where the propositional attitudes said to be known do not themselves exist. Thus, he criticizes as in bad faith those who avoid facing their knowledge of their absolute self-determining freedom. Yet most people do not actually have beliefs that they are totally free, beliefs that they can then in bad faith ignore; at best (and even this is doubtful), they have only 'implicit beliefs,' which are not actual thinkings at all but only potential ones (e.g., propositions easily inferable from actual beliefs and memories).

dition supposedly guaranteed by “the reflexivity of consciousness” (1998, 73). This awareness is what makes it true that “self-deception involves a pragmatic paradox in one’s relation to oneself,” a “double bind” (74). The self-deceiver’s “psychic system... oscillates indefinitely in the manner of the liar’s paradox” (81).

Anspach’s premise of self-awareness is itself suspect, and founded upon dubious Sartrean positions that will not be rebutted here. Moreover, the supposed self-awareness, as exemplified by the statements of a French proto-psychiatrist’s mental patient, does not show the paradoxical oscillations that Anspach supposes.

The alleged sufferer from madness (S) is quoted as saying

<1> If I could believe with you that I am mad, I would soon be cured, but I cannot acquire this belief. (79)

<2> ...I follow quite well your reasoning,... and if I could convince myself, I would no longer be crazy, you would have cured me. (82)

On Anspach’s reading, “on the one hand, he believes he is crazy, and yet... he does not believe he is crazy” (79). Note first that in stating that the psychotic both *believes and does not believe* he is crazy, Anspach makes what Davidson considers a common, but inexcusable, error. To Davidson (1986, 79–80; 1998, 5), such talk asserts ‘p and not-p,’ and so cannot make sense.

Yet it seems to be Davidson who is in the wrong here, for he discounts an implication of partitioning accounts. These allow that a person might in part of his mind believe that p, while in another part believe that not-p; so it should be possible in one part to believe p, while in the other to have no belief regarding p. If the parts are to some degree self-contained and independently usable thinking systems, this picture seems entirely reasonable. Then, just as ‘X believes that p, and X believes that not-p’—which, on its face, looks nonsensical given the ordinary notion of person-level belief; the evidence supporting each of the conjoined ascriptions undermining the other—actually makes sense, due to tacit semantic subtleties deriving from the background picture of mind, so might ‘X believes that p, and X does not believe that p’ make sense in similar fashion. In minds that are partitioned into sequentially activated, and sometimes inconsistent, perspectives, ‘believes and does not believe that p’ ascriptions may not be simply self-negations (i.e., ‘X believes that p, and it is not the

case that X believes that p’) or colloquial ways to express ‘believes that p and believes that not-p.’

That is, an excessive concern to avoid seemingly self-contradictory attributions may be inhibiting theorizing in this area. A further consequence of this is neglect of the reasonable possibility that some self-deception might involve driving oneself to or from a state of *non*-belief. If self-deception is an escape from an uncomfortable state of belief into a more tolerable one, then might it not sometimes involve fleeing from a disturbing state of undecided non-belief (about whether God exists, say, or whether one’s parent really loves one) into a state of belief, or from a state of belief into a state of nonbelief?

In any event, it is dubious to impute to S a belief that he is crazy. He never directly expresses such a belief, of course. To Anspach (1998, 79), however, S’s “I would /soon be cured/no longer be crazy” in <1, 2> implies the belief ‘I am not yet cured’; and then, “to believe he is not yet cured is to believe he is still crazy.” Hence, paraphrase <3>:

<3> *I am crazy*, and would cease to be if I were to come to believe that I am—but, I do not believe that I am.

If a fair paraphrase, this would not serve Anspach, because it would be an instance of Moore’s Paradox, which generally is thought to be absurd and not believable at all, rather than to involve oscillatory beliefs. Yet the best reading of <1, 2>, given a proper reluctance to attribute paradoxical belief or belief in contradiction, finds the more strongly counterfactual

<4> *If I were crazy*, I would cease to be, if I were to come to believe that am—but, I do not believe that I am.

S may, as Anspach states (79), “believe[] he *should* believe he is crazy” (and even wish he *could*), at least while conversing with his therapist, because he comprehends the latter’s reasoning and finds it intellectually appealing. But S cannot be said to believe he is crazy (as well as not crazy). This is no “mystery” (79)—it is commonplace to evaluate an entertained idea without finding grounds for objection, even while being precluded from also *believing* it by a subconscious recognition of conflicts between it and unspecified cognitions within one’s perspective (see Dennett 1978, 308).

Even were there evidence of the overt contrary beliefs 'I am crazy' + 'I am not crazy,' this would indicate *ambivalence*, not self-deception or paradox. It might also be straightforwardly explicable via partitioning, the two beliefs being available and supported in different regions of mind. Vacillation could occur, as Anspach notes—but due merely to alternating activations of the different parts, and in a way likely not to be "perpetual" (83) (as a change in either part, perhaps spurred by the discomfort of vacillating, could end the divergence in beliefs). (The same analysis holds if S is taken to have, not two contrary beliefs, but one belief ('I am crazy') and one state of nonbelief.)

104

Self-awareness of ambivalence ('I seem both to believe that p and to believe that not-p') does pose a further puzzle: how does partitioning provide a habitat for contrary beliefs if during reflection the segregation can be undone enough to allow such self-description? (The contradictory propositions still could not be used simultaneously in reasoning.)

The (non-Davidsonian) kind of partitioning mentioned above, with the mind activated one perspective at a time, ultimately may offer an answer. If one thinks with different networks of intentional states in different contexts of thought and action, one might sometimes accept that p, at other times accept that not-p, and in either type of context perhaps access a memory of the other. So when pondering an issue, one might shift back and forth between perspectives and consider the issue in their different lights. If no reconciliation of the standpoints is apparent, one may simply remain knowingly ambivalent for the time being.

But whether this story is convincing or not, the self-awareness of ambivalence does not entail paradox. Even an ensuing belief in one's own self-deception or madness—'I'm crazy because I believe that p and believe that not-p; yet I'm not crazy'—would only return one to (self-aware) ambivalence, though now with the different routes to judgment giving conflicting beliefs about *oneself*.

Real paradox requires the proper logical interlocking of propositions. How about (see Anspach 1998, 82)

<6> 'I (believe I) am crazy *because* I believe I'm not crazy'

<6'> 'I (believe I) am crazy *because* I don't believe I am crazy'

Suppose there were clear evidence of agents firmly believing (and not just

entertaining) <6> or <6'> due to a commitment to the (false¹⁴) theory expressed by the connective 'because.' Still, <6, 6'> must be construed as nonparadoxical, the second 'believe' referring to periods of time different from the first (e.g., ~~times other than this very moment of utterance~~) or to a nonconscious belief running in parallel in some partitioned-off part of mind. Here is where Davidson's admonition against ascribing belief in plain contradiction has its bite: the idea of having *conscious* occurrent beliefs 'p' and 'not-p' simultaneously just makes no sense.

This is not to say that no one can ever have oscillating beliefs. For example, one who believes he's not crazy, realizes what he believes, and holds 'Whoever does not believe he's crazy, is crazy (and inversely),' may come to believe he's crazy, then believe he's not crazy, and so on (though only until his concern or attention withers). What of it? At any time, he will have at most one of the beliefs—not contradictory beliefs contained in some special mental structure. Oscillation between two individually coherent beliefs is not tantamount to believing a paradox, which is incoherent and cannot be believed.

Belief in one's own self-deception is evidently not paradoxical or self-imprisoning, no matter how many rungs of self-mistrust one climbs. For example, I may believe I am self-deceived about my developing baldness, yet come to think that that second-order belief (about the corrupt nature of my first-order baldness-related belief) is itself corrupted, underlain by another, hidden second-order awareness that the first-order one is *not* a product of self-deception. This may seem perverse, but it might occur.¹⁵ The upshot, however, will tend to be the annihilation of my original concerns about the legitimacy of my baldness beliefs. There need be no oscillation—merely the conviction that I am not, after all, self-deceived about my hair, but am self-deceived about the epistemic status of my hair-belief. (See Appendix B for greater detail.)

14. (Dis)belief in one's sanity is no indicator of mental health: sane people, too, believe they are not crazy, and both sane and crazy people believe they *are* crazy. Further, while belief in one's insanity can increase receptiveness to treatment, it can also itself play a causal role in insanity. (These matters, however, are relevant only to madness and its cure, and not to self-deception and whether it involves paradox.)

15. Maybe facing the belief that I am not self-deceived about my baldness would require recognizing that I may lose the caring attentions of my therapist. (The newly uncovered second-order self-deception fails to strike me as likely to elicit her care.)

5. Lying to Oneself, the Liar's Paradox, and Double Binds

We have not yet encountered paradoxical beliefs that are 'möbiused,' rather than partitioned, within the mind. Anspach ultimately makes recourse to lying to oneself for the postulated element of paradox, but this tactic, too, is unavailing.

For even supposing that lying to oneself is present in self-deception, the idea that one who lies to himself is confronted with the liar's paradox (Anspach 1998, 80) is based upon a confusion. The liar's paradox applies only to reflexive assertions like 'I am *now* lying,' or 'This sentence is false'—and obviously not to an "I am lying" or "This statement is false!" remark (Anspach's example) that refers only to assertions other than itself. In a lie to others, such a disclosure is taken as a truthful warning, challenge, or confession, or perhaps a false joke or bluff. A hearer might reasonably scratch his head over the utterance's intended purpose, but not over whether its substance is self-referential and hence paradoxical. And it is clearer still that if an 'I am lying' belief is part of lying to oneself, it applies to the particular falsehood to be swallowed (e.g., 'I am not bald') and not also to itself. It is thus simply false that "the person on the receiving end [of a disclosure of an intention to deceive] is confronted with the liar's paradox" (Anspach 1998, 80).

Similarly, a psychotic's 'I am crazy' (83) need not itself be infected with madness. Anspach himself emphasizes that the mad are partly sane: "the most deluded lunatic can be quite rational" (63). If 'I am crazy' is a reflection upon other thought or behavior and not self-referential, talk of paradox is inapt.¹⁶ Not even 'All my beliefs, including *this* one, are delusions' would be paradoxical, for it could simply be that this belief is delusional (i.e., false) while some other(s) are not. (Compare: 'I always lie' can simply be false.)

It takes something quite peculiar to effect a liar's-like paradox. S asserts nothing of the sort. While the words 'I cannot acquire this belief' appear in <I>, "this belief" there refers not to those five words (or to <I>), but to another belief that <I> is about ('I am mad'). And if S's assertion had been

16. Anspach does take himself to have established that an 'I am crazy' judgment "applies not to any particular deluded belief external to itself, but to the very consciousness doing the judging," and that because it imputes "craziness as an overall outlook" it ought to be rephrased "This statement is untrue" (1998, 84, 84 n.12). His justification for this position is not at all clear.

self-referential? No matter. Liar-type propositions, though utterable as sentences to create paradoxes for interpreters, cannot be *believed* by those who fully understand them. One believes a proposition only when one holds it to be true, and if its content states that it is false, there is a psychological impossibility, and neither paradox nor oscillation. The 'if it's true, it's false, if it's false, it's true' nature of a liar's paradox has no clear relevance to lying to oneself.

Anspach, perhaps appreciating the force of this objection, has supplemented his discussion with remarks on Gregory Bateson's 'double bind,' which supposedly presents "an actual pragmatic dilemma" unlike the "purely semantic" liar's paradox (Anspach 1998, 74).¹⁷ He tells (74–76) of Bateson's example of a child and its unconsciously hostile mother, both of whom want to believe that the mother loves the child unreservedly. (The same dynamic can also exist within other relationships, and does not require hostility but simply a degree of feeling not as great as both wish to believe.) But Anspach never establishes his claim that (mutually dependent and mutually threatening) self-deception, rather than innocent, unmotivated misinterpretation is at work.¹⁸ More important, he never does

17. Anspach's examples arguably present no such dilemma. "Don't always do what I say" can be obeyed, while not thereby disobeyed as well: one need only disobey some future order out of compliance with this one. "Love me because you want to [or, Do D out of motive m], and not because I tell you to" can be obeyed: subsequent *D-ing-out-of-m* may be due to compliance with the order, but *D-ing* will even then be due not to compliance but to m, and hence will not constitute disobedience.

Other orders will naturally be construed as applying only henceforth, thus not to themselves (e.g., "Be spontaneous," "Think for yourself"). Even on reflexive construals, the results may again be only "semantic," descriptive paradoxes for observers, not actual pragmatic dilemmas for actors. "Disobey this order" is absurd, and affords no true alternative responses. As for "Disobey all orders, starting with this one," while obeying it does involve disobeying it, the converse can be escaped. If one chooses simply to disobey that order (by obeying some subsequent order), one has not thereby also obeyed it—though an observer might describe one as *in effect* having done so—because obedience to an order requires a motivation to comply with it, and that is, by hypothesis, absent. Receiving such an order, then, one faces no actual practical dilemma in forming an intention to act.

18. That is, it is not clear that either party ever realizes, in order to provide the initial motivation for self-deception, that the mother feels hostile toward the child (or that the evidence suggests that she so feels), despite Anspach's claims (76) that both "must" be aware of the truth; nor is it clear that the subsequent process of judgment is corrupt. And if they were truly self-deceived, would we not expect an end to the oscillation, as they would contrive a nonthreatening construal for the mother's periodic withdrawals?

show that in self-deception (and lying to oneself) one is caught in a double bind (76, 81).

Indeed, for all the eye-catching superficial resemblances among the various phenomena in Anspach's curiosity shop, it is not apparent that any of them really casts useful light upon self-deception.

Appendix A: 'Negative Cases'

108

One way in which Davidson distinguishes wishful thinking from self-deception is that supposedly in the former "belief takes the direction of positive affect, never of negative; the caused belief is always welcome," while "[t]he thought bred by self-deception may be painful" (1986, 87). Further distinctions, however, can clarify the contrast. The scale of absolute affect, from pleasurable to painful, ought not be confused with that of relative preferableness of states of affairs (or events) in the world, or of relative preferableness of states of belief.

Believing that is painful in an absolute sense—that 45 people have perished in an accident, that a lover's rejection was callous and selfish—can still be wishful thinking precipitated by "a desire or wish that a proposition be true" (Davidson 1998, 8)—but only if some more-dreaded scenarios have also been contemplated (as well, perhaps, as none that seems substantially preferable and reasonably likely). If the alternative is that 150 have died, or that the lover's rejection was honorable and caring, thus making the loss more unbearable, the thinking may be wishful despite being distressing.

Jealously believing that one's spouse had an illicit sexual encounter during a business trip thus may be wishful thinking, if one has contemplated a worse alternative state of the world that would explain one's evidence (say, that the spouse has been conducting a long-term affair)—but otherwise is not, even if one *has* sensed that another state of belief would be less preferable (e.g., if one feels relief at not taking one's spouse to be faithful and thereby leaving oneself maximally vulnerable to devastating later disappointment and humiliation.) So, in this case, the jealous belief may perhaps be innocent, or might involve some kind of irrationality such as self-deception or fretful thinking (see below); but it is not a wishful thinking.

Corresponding observations may apply, with the polarity of terms reversed, to 'fretful thinking,' a counterpart phenomenon that seems rec-

ognized by folk psychology despite not being named. Just as some beliefs are taken to result directly from *wishes* that certain propositions be true, and not indirectly (and self-deceptively) from aversion to other existing opposed beliefs, so do some beliefs seem to be produced directly (without rational evaluation of evidence, and without self-deception) by *worries* or *fears* that certain propositions are true. Though these new beliefs depict relatively undesirable states of the world, they may or may not be pleasurable in an absolute sense. In either case, by hypothesis they are not desired states of belief—hence the explanation for their production cannot be intentional self-deception or nonintentional 'wishful belief' (Talbot 1995, 59).

If this is right, it would then seem not to be the case that, "If a pessimist is someone who takes a darker view of matters than his evidence justifies, every pessimist is to some extent self-deceived" (Davidson 1986, 87). For pessimism might be produced by fretful thinking alone.

Appendix B: Believing in One's Own Self-Deception

Does believing that one is self-deceived plunge one into paradox? Consider the example in the text (pp. 104–105). At the outset, my relevant belief is

<a> I am not growing bald.

To which gets added

 I am self-deceived; <a> is the product of self-deception.

From it follows that

<c> I *really* believe/know that I *am* growing bald.

So when is in mind, I also employ as a belief the proposition

<a'> I am growing bald.

At this stage, there is no paradoxical indeterminacy. Speaking loosely, there may be an 'oscillation' of sorts, because whether I use <a> or <a'>

will vary with the perspective then in place. If is not in mind at the time, I will freely use <a>. If is present, I will use the epistemically untainted <a'>, however much <a'>'s intuitive appeal. (Of course, if the appeal grows so strong that I forget about for the time being, I will revert to <a>.) Thus it seems quite possible to acknowledge one's self-deception without remaining mired in it.

The next, more perverse stage of self-doubt adds:

<d> I am self-deceived regarding ; i.e., is the product of self-deception.

From which it follows that

<e> I *really* believe/know that I am *not* self-deceived regarding <a>, i.e. that <a> is *not* the product of self-deception.

So when <d> is in mind, I also employ as a belief the proposition

<b'> I am not self-deceived regarding <a>; i.e., <a> is not the product of self-deception.

With an added layer of self-skepticism, then, there is again a non-paradoxical variation in belief state. If the perspective of the moment includes <a> but neither nor <d>, I will use the proposition <a> in thought. If it includes but not <d>, I will use <a'> as discussed above. If, however, it includes <d>, I will use <a>, because <d> undermines the perceived validity of while has no reciprocal impact upon <d>.

References

- Anspach, M.R. 1998. "Madness and the Divided Self: Esquirol, Sartre, Bateson." Paper presented at the Symposium on Self-Deception, Stanford, California, February 1993. In J. P. Dupuy, ed., *Self-Deception and Paradoxes of Rationality*. Stanford: CSLI Publications (this volume).
- Davidson, D. 1982. "Paradoxes of Irrationality." In R.A. Wollheim and J. Hopkins, eds *Philosophical Essays on Freud*. Cambridge: Cambridge University Press.
- . 1985. "Incoherence and Irrationality." *Dialectica* 39:345–54.

- . 1986. "Deception and Division." In Elster, ed. (1986)
- . 1998. "Who Is Fooled?" Paper presented at the Symposium on Self-Deception, Stanford, California, February 1993. In J. P. Dupuy, ed., *Self-Deception and Paradoxes of Rationality*. Stanford: CSLI Publications (this volume).
- Dennett, D. C. 1978. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, Mass.: MIT Press.
- Elster, J., ed. 1986. *The Multiple Self*. Cambridge: Cambridge University Press.
- Heil, J. 1994. "Going to Pieces." In G. Graham and G. L. Stephens, eds., *Philosophical Psychopathology*. Cambridge, Mass.: MIT Press.
- Mele, A. R. 1987. *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. New York: Oxford University Press.
- . 1998. "Two Paradoxes of Self-Deception." Paper presented at the Symposium on Self-Deception, Stanford, California, February 1993. In J. P. Dupuy, ed., *Self-Deception and Paradoxes of Rationality*. Stanford: CSLI Publications (this volume).
- Nisbett, R. and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Pears, D. 1986. "The Goals and Strategies of Self-Deception." In Elster, ed. (1986).
- Sartre, J.-P. 1956. *Being and Nothingness*. Trans. H. Barnes. New York: Philosophical Library.
- Talbott, W. J. 1995. "Intentional Self-Deception in a Single Coherent Self." *Philosophy and Phenomenological Research* 55:27–74.
- Wood, A. W. 1988. "Self-Deception and Bad Faith." In B. P. McLaughlin and A. O. Rorty, eds., *Perspectives on Self-Deception*. Berkeley: University of California Press.