

What's Bad About Bad Faith?

Our contemporary common sense ethical framework is committed to an ideal of authenticity, according to which there is something bad about being inauthentic, and something good about authenticity. Call this the BASIC INTUITION about authenticity. There is a related, but distinct facet of the ideal of authenticity: we are critical of the inauthentic person, such that to call someone a phony, a fake, or a hypocrite is to accuse her of something, to offer some kind of censure of her or of her behavior, and to charge her with doing or having done something wrong. And we praise authenticity even when manifested in behavior that we would otherwise condemn; someone's bad behavior is vindicated for us on the grounds that, at least, it was authentic, that the vicious actor was "being herself." The common sense ideal of authenticity, therefore, also says that the inauthentic person is doing something wrong and is culpable for her inauthenticity. This is why the characterization of another as inauthentic takes the form of a charge or an accusation. Call this the CULPABILITY INTUITION about inauthenticity.

Supposing the basic intuition is right, what explains its truth? In what sense, exactly, is inauthenticity bad? On standard philosophical accounts of authenticity, inauthenticity is bad because of what the inauthentic person does *to herself*. Charles Taylor writes that (according to the common sense ideal of authenticity) being authentic is a "better or higher mode of life" (1991, p. 16) than is being inauthentic, and that to fail to live up to this ideal is to live a life that is "somehow wasted or unfulfilled." (p. 17) Developing his own account of inauthenticity, or what he calls "bad faith," Sartre maintains that bad faith is a kind of self-deception: the inauthentic person does herself harm by lying to herself. (1956, p. 86-96) The common commitment here is that inauthenticity is bad *for the inauthentic person*, in the sense that inauthenticity is a threat to the goodness or quality of a person's life. We have in mind here the sense of quality of goodness evoked by Parfit's question:

What would be best for someone, or would be most in this person's interests, or would make this person's life go, for him, as well as possible? (1984, p. 493)

So we shall understand this standard picture of authenticity to essentially involve a commitment to the following account of the badness of inauthenticity, affirming a connection between authenticity and personal wellbeing:

WELLBEING ACCOUNT: No good life involves substantial, prolonged, or systematic inauthenticity.

In this paper we criticize three versions of the wellbeing account, and propose an alternative account of the badness of inauthenticity. In §1, we consider the metaphysically suspicious claim that authenticity requires fidelity to one's "true self." In §2, we consider an intuitive account of authenticity as not pretending to be other than you are, but conclude that there can be good lives that involve substantial, prolonged, or systematic pretending, and thus that the wellbeing account fails, on this conception. In §3, we consider the Sartrean idea that inauthenticity is a kind of self-deception, and again conclude there can be good lives that involve substantial, prolonged, or systematic self-deception, and thus that the wellbeing account fails. Finally, in §4, we reject the wellbeing account, and propose instead that

inauthenticity is bad because the inauthentic person runs the risk of deceiving other people. On our diagnosis, the badness of inauthenticity is explained by the fact that the inauthentic person always runs the risk of treating others in morally problematic ways. And because we conceive of the badness of inauthenticity in moral terms, our account, unlike the wellbeing account, makes sense of the culpability intuition.

1.

What is it for a person to be inauthentic? Consider what seems like a paradigm case of inauthenticity. Smith has always been a passionate baseball fan, having spent many Sunday afternoons at Fenway Park with her family as a child. She regularly attends games, wears Red Sox t-shirts, orders special baseball cable channels, and so on. But when Smith moves from Providence to Boston, to work at a new job, she finds that her new colleagues don't share her interest in baseball. They are, however, obsessed with basketball, and they repeatedly invite her to join them at games, to go out to watch the games on television, to talk about basketball at lunch, and so on. Smith, wanting to fit in with her new colleagues, decides to fake it: she buys some Celtics gear and joins her colleagues, pretending to like basketball, and her relationships with her new friends flourish.

What makes this a paradigm case of inauthenticity? A natural way to articulate the notion of authenticity is to say that authenticity is being true to yourself, so that being inauthentic involves some kind of betrayal of oneself. But what do *these* notions amount to? What does it mean for a person to be true, or false, to herself?

One kind of answer to this question is offered by Taylor:

Herder put forward the idea that each of us has an original way of being human. Each person has his or her own "measure" is how way of putting it. [...] There is a certain way of being human that is *my* way. I am called upon to live my life in this way, and not in imitation of anyone else's. [T]his gives new importance to being true to myself. If I am not, I miss the point of my life, I miss what being human is for *me*. (1991, pp. 28-9)

In the language of contemporary popular psychology, each person has a "true self," which is something (or someone?) distinct from her. A person's behavior may or may not manifest her "true self," but her life will be "somehow wasted or unfulfilled" (as Taylor puts it) unless she is faithful to it. Taylor goes on to speak of the importance of "a kind of contact with myself, with my own inner nature." (p. 29) So authenticity amounts to fidelity to my "inner nature" or "true self."

But there are decisive metaphysical problems with this account. If people have "true selves" or "inner natures," these can't be capable of changing. Suppose Smith's story ends like this: After a few months of watching basketball games with her new friends and after her friends explain a few things about the game, Smith finds herself enjoying it, and appreciating aspects of basketball she hadn't understood before; eventually, she becomes as genuinely passionate about basketball as any life-long Celtics fan. Now the natural thing to say about this is *not* that Smith "betrayed her true self" (who loved baseball), nor is it to say that she finally "found her true self" (who loves basketball). The natural thing to say is that *Smith changed*.

The defender of the idea of an “inner nature” or “true self” faces the question: which of Smith’s selves was her “true self”? And what makes that version of Smith the “true” one, and the other the “false”?

It’s not plausible to say that the *source* of Smith’s interest in basketball is the culprit. If we change the story so that Smith’s father is responsible for her previous love of baseball (she only went to the games to please him, but eventually found herself enjoying them), shall we then say that she was betraying her “true self” even before her meeting her new friends? And if an origin like that is sufficient for some aspect of a person not being part of her “true self,” what elements of our identities could be said to be authentic? We develop our loves and hates, our values, our interests, and everything important about us as persons, through interactions with what is “external” to us: other people, social institutions, our physiology. So an “external” origin can’t explain why Smith’s concern for basketball isn’t part of her “true self.” And, in general, being changed and influenced by “the demands and expectations of others” is a normal and unobjectionable part of most human lives.

For these reasons, we are reluctant to take talk of “true selves” and “inner natures” seriously. If the talk of “inner natures” and “true selves” is meant as metaphor, that is fine, but the wellbeing conception can be made plausible only if we can give something like a non-metaphorical articulation of what being authentic amounts to.

2.

Although it is obscure what it would mean to say that Smith’s “true self” prefers baseball to basketball, there is a non-obscure truth in the neighborhood, namely, that *Smith* prefers baseball to basketball. In what way does she betray herself, then, by pretending to like basketball? Perhaps her inauthenticity amounts to nothing more than that: to the fact that she is pretending. In the passage from Taylor, quoted above, we focused on the appeal to “inner natures.” But perhaps the emphasis should have been on the idea that authenticity essentially involves avoiding the imitation of other people. Consider the following “paradigm case of inauthenticity,” from David Velleman:

This person laughs at what he thinks he is supposed to find amusing, shows concern for what he thinks he is supposed to care about, and in general conforms himself to the demands and expectations of others. The motives that his behavior is designed to simulate are motives that he doesn’t genuinely have. (2002, p. 97)

As Lionel Trilling puts it, authenticity requires “a congruence between avowal and actual feeling.” (1972, p. 2, see also p. 27) Inauthenticity is pretending to be someone or something other than you actually are; authenticity is not pretending. Is the wellbeing account of the badness of inauthenticity plausible, on this conception of authenticity?

We think not. Consider Jones, a detective playing the “bad cop” role in an interrogation. Jones glares angrily at the suspect, she slaps a cigarette out of his mouth, she declares that she “doesn’t have time for scum like you,” and storms out of the room. The motives that Jones’ behavior is designed to simulate are not motives that Jones genuinely has; Jones

suffers from a lack of congruence between avowal and actual feeling. She is pretending to be other than she is, and thus is inauthentic, with respect of the interrogation.

Jones seems no worse off for her inauthenticity. (Or, at least, we can easily imagine that she is no worse off: if she enjoys playing the “bad cop,” or if she just doesn’t care about it and sees it as part of her job.) But our question is whether there might be good lives that involve substantial, prolonged, or systematic inauthenticity, which on the present conception is identified with pretending, so we must ask: are there good lives that involve substantial, prolonged, or systematic pretending?

A negative answer to this question would entail that certain careers are incompatible with living a good life: being an undercover police officer, or a politician, or an actor, or a café waiter. All these jobs involve a great deal of pretending: pretending to be someone else, pretending to have feelings you don’t have, pretending to be happy when you aren’t, and so on. It would be a mistake to insist a priori that the life of someone who works one of these jobs can’t be a good life, just in virtue of being a job that necessarily involves (substantial, prolonged, or systematic) pretending.¹ Substantial pretending is a normal and unobjectionable part of many human lives.

This doesn’t mean that there aren’t cases in which pretending significantly detracts from the goodness of a person’s life. Consider Jews passing as Christians during the Inquisition or the Holocaust. It could be maintained that no good life involves *this* kind of pretending; that these Jews were deprived of the chance of living a good life, by being forced or coerced into pretending to be Christians. This kind of pretending might also detract from the goodness of a person’s life in virtue of a powerful and unpleasant feeling of alienation that might go with it.² Our point is that (substantial, prolonged, or systematic) pretending, *as such*, is not incompatible with living a good life. For someone coerced by threat of violence into pretending not to be Jewish, what threatens the goodness of her life is not her inauthenticity, but that fact that her inauthenticity and the accompanying feeling of alienation was the result of being coerced by threat of violence. And being coerced by threat of violence to do *anything* is a threat to the goodness of a person’s life.³

Smith’s inauthenticity seems to threaten the goodness of her life precisely if we add details to her story that make it look more like a case of forced pretending. If she feels powerless to resist the “peer pressure” to become a basketball fan, or if she finds the prospect of not getting on with her colleagues so terrifying that she reluctantly accepts the only slightly-less horrible option of pretending to enjoy basketball, then her life is not going well. But, as

¹ Our argument doesn’t depend on the idea that pretending is required for many (wage-earning) jobs. Someone could pursue acting as a creative or artistic project, not as a means to earn a living, but as a meaningful ground project. The claim that such a person could not live a good life is implausible.

² Given this, we can distinguish between being alienated from the pretending involved in one’s job (which plausibly threatens the goodness of one’s life) and simply pretending, as part of one’s job (which, we maintain, does not necessary threaten the goodness of one’s life).

³ [Reference to authors’ paper.]

above, in that case, the threat to Smith's wellbeing is not coming from her inauthenticity, as such, but from the fact that she is being forced to be inauthentic.

On the other hand, Smith's inauthenticity seems not to threaten the goodness of her life at all if we add details that reduce or eliminate the sense of coercion. If she consciously recognizes that she prefers friendship to friendless baseball fandom and deliberately sets out to make friends by feigning an interest in basketball; if she thinks of her pretending as an opportunity for desired change; or if she is open to learning to eventually like basketball, and so on, then her inauthenticity does not detract from the goodness of her life. Indeed, pretending may be an essential part of the life-expanding experimentation that is characteristic of the very best lives. It seems virtuous to try things we are unsure about (or even downright suspicious of). And when we do, we will often commit to them in our avowals or our behavior, if not (yet) in our feelings. We call this "going for it" and we praise people for their open-mindedness and their adventurous spirit when they do.⁴

Pretending may also be an unavoidable part of becoming who we are. When a college freshman changes her habits of dress in the process of finding a comfortable social place, her high school friends may accuse her of being a phony (and she may even feel as if she is faking it); but we can far more charitably describe her as growing up, as figuring out who she is (as believers in "true-self" metaphysics might put it), or simply as changing over time. Such change, resulting from experimentation, is familiar, typical, and unobjectionable.⁵

Given all this, we conclude that the wellbeing account of the badness of inauthenticity fails, if we conceive of authenticity as not pretending. We sympathetically return to this conception of authenticity (but not to the wellbeing account) in §4.

3.

We turn now to Sartre's account of inauthenticity. In §2, being inauthentic was conceived of as a way of acting or behaving; on Sartre's conception, being inauthentic is a way of thinking, and in particular a way of thinking about oneself. In a well-known passage, Sartre describes what he takes to be a paradigm case of "bad faith": the case of a woman whom we'll call Michelle. Michelle "has consented to go out with a particular man for the first time."

She knows very well the intentions which the man ... cherishes regarding her. She knows also that it will be necessary sooner or later for her to make a decision. But she does not want to realize the urgency; she concerns herself only with what is respectful and discreet in the attitude of her companion. [...] But then suppose he takes her hand. This act of her companion risks

⁴ See §4 where we offer a related comment on Sartre's ambivalent young woman with her date. If she is required to be wholehearted (or at least to behave as if she is), the call to authenticity may itself turn out to require a certain kind of pretending.

⁵ Marcia Baron says something similar about self-deception: it is often best to "stumble along confusedly ... settling on *some* plan of action and *some* picture or other of what is going on, in an attempt to end the painful confusion and unease and get on with life. In fact, it is not that easy to figure out what we are about, discovering it may at times have to arise *through* self-deception." (1988, p. 441)

changing the situation by calling for an immediate decision. To leave her hand there is to consent in herself to flirt, to engage herself. To withdraw it is to break the troubled and unstable harmony which gives the hour its charm. The aim is to postpone the moment of decision as long as possible. We know what happens next; the young woman leaves her hand there, but she *does not notice* that she is leaving it. (1956, p. 96-7)

Sartre maintains that Michelle engages here in a particular kind of self-deception. (In §4 we shall offer an alternative account of what Michelle's inauthenticity consists in.) He develops this account in the context of a philosophical framework with which we have substantial sympathy. Sartre rejects the idea of a "true self" and of an antecedently given "unique way of being human" for each of us, and recognizes the difficulty in giving a metaphysically adequate account of what it would be for a person to be authentic or inauthentic. He proposes that inauthenticity amounts to self-deception about one's own human condition. He says, of Michelle:

[W]hile sensing profoundly the presence of her own body ... she realizes herself as *not being* her own body, and she contemplates it as though from above as a passive object to which events can *happen* but which can neither provoke them nor avoid them[.] The basic concept which is thus engendered utilizes the double property of the human being, who is at once a *facticity* and a *transcendence*. These two aspects of human reality are and ought to be capable of a valid coordination. But bad faith does not wish either to coordinate them or to surmount them in a synthesis. (p. 98)

Michelle's inauthenticity amounts to her refusal to recognize her own status as a free agent (a "transcendence"); she treats herself as though she were merely an inanimate, unthinking object (a "facticity"), thereby refusing to face up to the necessity of making a choice about whether to respond positively or negatively to her date's advances. Inauthenticity is a "lie to oneself" (p. 89) about one's own freedom; it involves "hiding something from oneself" (p. 94), and it is one's own freedom that one hides from. Summing up this existentialist account of inauthenticity, Simone de Beauvoir writes that the person in bad faith "cannot efface the agonizing evidence of his freedom." (1948, p. 45) Authenticity, then, would amount to avoiding this kind of self-deception, and wholeheartedly accepting one's status as both a facticity and a transcendence.

Is this plausible, as a version of the wellbeing account of the badness of inauthenticity? In other words, is it plausible to maintain that no good life involves substantial, prolonged, or systematic inauthenticity, when inauthenticity is understood along these lines? We have two reasons for giving a negative answer to this question.

The first is that Sartre's account essentially relies on his controversial picture of the human condition: his account is correct only if it is true that human beings possess the kind of radical freedom they are said to possess by existentialist philosophers. As a result, the wellbeing conception of the badness of inauthenticity, combined with Sartre's conception of inauthenticity, has the consequence that no one can live a good life unless they agree with Sartre about the metaphysics of personal freedom. We're sympathetic with Sartre's views, but refusing to accept one's status as a transcendence is a philosophical mistake, and

moreover it doesn't seem like the kind of philosophical mistake the making of which would preclude someone from living a good life.

Sartre's critique of inauthenticity falls squarely within the traditional view in modern psychology and sociology that, as Taylor puts it, "the ideally strong character ... would be able to face unflinchingly the truth about himself or herself." (1989, p. 33) A Sartrean wellbeing account of the badness of inauthenticity requires something like this claim: that no one lives a good life if she lacks self-knowledge. Our second reason for rejecting such an account is that we think this claim should be rejected. This is because, in our view, substantial and systematic self-deception is the norm for most people, and, far from threatening the goodness of their lives, said self-deception makes a positive contribution to their ability to live well. The view we have in mind has its origins in Hume, who noted the "almost universal propensity of men, to over-value themselves." (*Treatise*, Book III, Part III, Section II) However, "a genuine and hearty pride, or self-esteem, if well conceal'd and well founded, is essential to the character of a man of honor." (Ibid.) Contemporary research by social psychologists has for the most part supported Hume's contention; what he called "self-applause" is now known as "self-enhancement bias," which is both widespread and which, more importantly, tends to be absent in depressed subjects. A number of studies have confirmed non-depressed subjects' tendencies to rate themselves as "above average" or "better than most," across a wide range of features and abilities: intelligence, leadership ability, competence at one's job, athletic ability, and so on. (See Taylor and Brown 1988 and Brown and Dutton 1995 for reviews of the literature.) Subsequent studies found that non-depressed subjects overestimated their positive qualities, by comparing their self-assessments with assessments by third-party observers (e.g. John and Robins 1994). For Hume, self-esteem is required to be a "man of honor"; for the social psychologists, it's required for (or at least correlated with) psychological health. Suppose that Hume and the psychologists are right, that irrational overestimation of one's positive qualities is both widespread and required for (or at least correlated with) living well. If this is right, and if the badness of inauthenticity is conceived of essentially as a failure of self-knowledge (as it seems to be on Sartre's conception), then the wellbeing account of the badness of inauthenticity is false.

Marcia Baron makes the same point, in a discussion of the badness of self-deception:

Consider some other cases: the AIDS victim who deceives himself into believing that his chances of recovery are at least fifty-fifty; the Salvadoran who convinces himself that his sister escaped into the mountains or fled to Mexico, and was not, despite the evidence, tortured and killed by the army's death squads. It is difficult to find these instances of self-deception objectionable. (1988, p. 433)

And Baron goes on to describe a plausible case of the self-enhancement bias at work:

[A] historian denied tenure at a second-rate school gets a job at a considerably inferior school. Imagine that he is good but not first-rate, and that he was judged by the tenure committee to be a far worse scholar and teacher than he is. In reaction to the tenure decision he exaggerates the gap between the assessment of him and what a fairer assessment would have

yielded. He compares his work to that of major contemporary historians and judges his to be as good or better than most of their work. (pp. 433-4)

On our view, the historian's self-deception does not necessarily threaten the goodness of his life. Even if he persists in his irrational belief (so that his self-deception is prolonged), or even if he is generally disposed to engage in irrational self-enhancement (so that his self-deception is systematic), his life is not *thereby* excluded from being a good life. After all, if Hume and the social psychologists are right, most people are generally disposed to engage in such thinking; most people do it all the time. We find it hard to believe that something so ubiquitous could preclude living well. Perfect self-knowledge may be some kind of ideal, but it is not a requirement on living well (at least not for people like us).⁶

Baron argues that self-deception is dangerous because it threatens to undermine our agency, just as deception of others, because it requires manipulation, threatens to undermine the agency of others. And since "self-deception can become a habit, a strategy one falls back on too often" (p. 437), and since "self-deception often requires, for its efficacy, further self-deception" (Ibid.), it is particularly dangerous. If self-deception undermines agency, and inauthenticity is a kind of self-deception, then might the wellbeing account be correct after all? Two replies to this.

First, it's not obvious that enjoying a lower degree of agency (as a result of one's having undermined one's own agency) necessarily detracts from the goodness of a person's life, such that no good lives involves such self-undermining. It indeed seems bad to be manipulated, when we consider cases in which someone is manipulated by another person. But it is bad to be manipulated by yourself? That seems to be the question we started with; so the appeal to the idea that self-deception undermines agency hasn't added anything.

Second, there are reasons to be skeptical of the idea that self-deception always threatens to "spread" in the way that Baron suggests. The thought seems to be that if a person believes *p* as a result of self-deception, she will be rationally required to believe various other things (e.g. things implied by *p*) or to revise beliefs of hers that conflict or fail to cohere with the belief that *p*, and thus the self-deception will "spread." But this assumes that the self-deceiver will do what is rationally required, once she acquires the belief that *p*. Why assume that someone who engages in an irrational strategy of belief formation (the initial self-deception) will subsequently transform into a rational believer? Indeed, two familiar aspects of beliefs based on this kind of irrationality are (1) that believers are remarkably resistant to evidence against the irrational belief, and (2) that irrational beliefs of this kind become "compartmentalized," so that believers fail to draw conclusions from them and to incorporate them into other parts of their system of beliefs. The amateur basketball player who deceives himself into thinking that he could be competitive in the NBA doesn't conclude from this that he should go to NBA tryouts; he may even know that he wouldn't

⁶ Sartre suggests that the threat of bad faith is an inescapable feature of the human condition (p. 116). If this is the case, and if the wellbeing account is right, Sartre is making a rather bleak observation about our prospects of living good lives. In our view, Sartre's proposal makes is an over-demanding requirement, for people like us.

be competitive in the NBA, and thus have conflicting beliefs, but his irrationality may prevent him from noticing this or giving it any attention.

So, to sum up: if we conceive of inauthenticity as Sartre does (as self-deception about the metaphysics of personal freedom), then the wellbeing account of the badness of inauthenticity fails, as (i) being mistaken about philosophy is not incompatible with living a good life, and (ii) self-deception, in general, is not incompatible with living a good life.

4.

Recall Michelle, who lets her date hold her hand while, in some sense, “not noticing” that she’s letting him. Note that the story requires that Michelle “does not quite know what she wants” (Sartre 1957, p. 97); this is what makes her reluctance to make a decision psychologically plausible. She can “not know what she wants” in either of two senses. On the one hand, she might simply be ignorant of certain facts which would settle things for her one way or the other: if he’s a Laker fan, then that’s a dealbreaker; she wholeheartedly wants to be romantically involved with someone who isn’t a Laker fan, and her “not knowing what she wants” just consists in her not knowing whether this particular man is a Laker fan or not. On the other hand, she might have all the relevant information, but remain conflicted or ambivalent about whether she wants to be involved with the man she’s on a date with. This will be the case if she has desires that conflict, either necessarily or because of contingent facts about this particular man: she’s physically attracted to him, but finds him dim-witted, for example. Her “not knowing what she wants” here consists in her having desires that conflict.

“Not knowing what you want,” in either of these senses, isn’t intuitively a form of inauthenticity.⁷ However, there *is* a sense in which Michelle is inauthentic in Sartre’s story, but it’s not a matter of her beliefs (e.g. her engaging in self-deception) or her desires (e.g. her “not knowing what she wants”). She is inauthentic when she doesn’t withdraw her hand, for in doing this she continues her flirtation, and (so Sartre seems to think, anyway) thereby gives her companion the impression that she *does* “know what she wants,” namely, to become romantically involved with him. In not withdrawing her hand, she gives the impression that she *has* made a decision, when in fact she has not. In acting as she does, she runs the risk of misleading or deceiving him.

There is something very strange about Sartre’s description of this case, which may get in the way of appreciating the point we’re trying to make. Michelle is described as being under pressure to “make a decision” about something, a “necessary” decision, but about what, exactly? Whether to have sex with her companion? Whether to make some kind of avowal or confession of attraction? Whether to let him hold her hand? The latter seems the only decision Michelle is forced to make in the course of the story as Sartre tells it, and thus the only decision that seems in any obvious way “necessary.” And what are the “intentions which the man ... cherishes regarding her”? The intention to have sex with Michelle? The intention to take her out again? The intention to marry her? How does Michelle “know very well” what these intentions are? If she does know very well what the intentions of her

⁷ There are, however, critics of ambivalence, e.g. Frankfurt 1992. [Reference to authors’ paper.]

date are, she'd be the first person to ever enjoy such certain knowledge of the mind of a prospective romantic partner. Why are her companion's intentions presumed to be determinate and clear, when Michelle's own intentions are indeterminate and unclear? It can be difficult to take Sartre's case seriously, once its implicit misogyny is on the table: men know what they want, they're clear about what they want, and everything would be fine if fickle women would just make up their minds.

However, our point is that Michelle does seem to run the risk of misleading or deceiving the man she's on a date with, if only because he unjustifiably expects some kind of assurance or "sign" that she feels a certain determinate way about him. Let's grant the dubious and problematic assumption that holding hands with someone *does* constitute an avowal of non-ambivalent romantic interest; there is, then, an incongruence between Michelle's avowal and her actual feelings, to borrow Trilling's phrase. And it is in virtue of this incongruence that Michelle runs the risk of misleading or deceiving her companion.

This suggests an alternative to the wellbeing account of the badness of authenticity. On this alternative account, which we favor, inauthenticity is bad because the inauthentic person always runs the risk of misleading or deceiving other people. To the extent that this is morally problematic, being inauthentic is morally problematic as well. On this view, inauthenticity isn't bad because of how it affects the authentic person herself, it's bad in virtue of how it affects (or threatens to affect) other people.

So we propose to revive the idea, considered in §2, that inauthenticity essentially requires acting in a way not representative of one's actual beliefs, desires, and other attitudes. But rather than combining this with a wellbeing account of the badness of inauthenticity, we propose a *moral account* of the badness of inauthenticity. There are two plausible versions of this account. According to the first:

MORAL ACCOUNT, VERSION 1: Being inauthentic is morally problematic when, and only when, it results in someone being misled or deceived, and it is problematic to the extent, and only to the extent, that someone is misled or deceived.

Someone deserves censure for being inauthentic only when said inauthenticity is morally problematic. Thus Smith's phoniness, her posing as a basketball fan, can warrant censure only if her colleagues are deceived into thinking that she is a basketball fan. If not (e.g. if she tells them about her indifference to basketball), then, on our view, there is nothing wrong with her inauthenticity.

There are reasons to prefer a less permissive version of the moral account, on which:

MORAL ACCOUNT, VERSION 2: Being inauthentic is morally problematic when, and only when, the inauthentic person runs a risk of misleading or deceiving someone, and on which it is problematic just to the extent that said risk is substantial.

Thus even if Smith's colleagues aren't actually misled, her inauthenticity will still be morally problematic if she runs some substantial risk of misleading or deceiving them (or others).

There are two reasons to favor this less permissive version of the moral account. First, given our conception of inauthentic behavior (as incongruence between action and real attitudes), inauthentic behavior is a kind of behavior that, in general, has the potential to mislead or deceive people. Since it is plausibly morally problematic to engage in a *kind* of behavior that has that potential, it is morally problematic to be inauthentic. Second, someone who engages in inauthentic behavior may become disposed to engage in it in the future. Thus, given the first reason, just stated, anyone who engages in inauthentic behavior runs the risk of inculcating a morally problematic habit.⁸

On this version of the moral account, a person might deserve censure even if her inauthenticity neither misleads nor deceives anyone: she deserves censure for engaging in the *kind* of behavior that can mislead and deceive people, and for engaging in behavior that might make her disposed to mislead or deceive people in the future.⁹

Two caveats. First, even if someone's inauthentic behavior is morally problematic (as it sometimes will be, according to both versions of the moral account), it may not be morally impermissible: consider the Jews who denied their Jewishness to escape the concentration camps. To say that some behavior is morally problematic is to say that one always has a *prima facie* (or perhaps even *pro tanto*) moral reason to avoid that behavior. Second, that someone's inauthentic behavior is morally problematic, even morally impermissible, does not entail that it would not be best for that person, all things considered, to be inauthentic. Consider Michelle. Perhaps misleading her date is morally problematic, perhaps nothing else speaks morally in favor of her misleading him, so that from the moral point of view, she ought not hold his hand. She might still be better off holding his hand: if she thinks withdrawing her hand will make him angry, or create a scene, or perhaps just because she enjoys ambiguously flirting with him, despite the fact that this involves misleading him. The moral account says that being inauthentic is sometimes morally bad; this doesn't mean that being inauthentic is so morally bad that good people don't sometimes indulge in it.

The second version of the moral account captures the basic intuition about authenticity, that there is something bad about being inauthentic, and something good about authenticity. In this respect, the moral account succeeds where the versions of the wellbeing account that we considered above fail. But the moral account also succeeds in vindicating the culpability intuition: because being inauthentic is often morally problematic, the inauthentic person will often deserve censure. (But, given the two caveats, above, there might be cases of morally problematic inauthenticity in which it is morally best, or all-things-considered best, not to censure the inauthentic person.) A person can be accused of doing something morally problematic, in a way that she cannot be accused of not living well. For this reason, the culpability intuition does not sit easily with the wellbeing account; but it is a natural corollary of the moral account.

⁸ The success of this argument for the second version of the moral account depends on an empirical assumption, common to many virtue theoretic accounts of morality: that engaging in a certain kind of behavior tends to habituate one to that kind of behavior. To the extent that this is false (and in particular, to the extent that it is false when it comes to inauthentic behavior), the first version of the moral account may be preferable to the second version.

⁹ This defense of the moral account mirrors some remarks Baron makes about the self-deceiver's capacity to mislead others; see her 1988, pp. 434-5.

Bibliography:

- Baron, M. (1988), "What is Wrong with Self-Deception?" in McLaughlin and Rorty (eds.), *Perspectives on Self-Deception* (University of California Press), pp. 431-49
- Beauvoir, S. de (1948), *The Ethics of Ambiguity* (Citadel Press).
- Brown, J.D., and Dutton, K.A. (1995), "Truth and Consequences: The Costs and Benefits of Accurate Self-Knowledge," *Personality and Social Psychology Bulletin* 21, 1288-1296
- Frankfurt, H. (1992), "The Faintest Passion," *Proceedings and Addresses of the American Philosophical Association* 66, reprinted in his *Necessity, Volition, and Love* (Cambridge University Press, 1999), pp. 95-107.
- John and Robins (1994), "Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism," *Journal of Personality and Social Psychology* 66, pp. 206-19
- Parfit, D. (1984), *Reasons and Persons* (Oxford University Press).
- Sartre, J. (1957), *Being and Nothingness* (Washington Square Press).
- Taylor, C. (1989), *Sources of the Self* (Harvard University Press).
- (1991), *The Ethics of Authenticity* (Harvard University Press).
- Taylor, S. and Brown, J.D. (1988), "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin* 103, 193-210
- Trilling, L. (1972), *Sincerity and Authenticity* (Harvard University Press).
- Velleman, J.D. (2002), "Identification and Identity," in Buss and Overton (eds.), *Contours of Agency* (MIT Press), pp. 91-123.