

Forthcoming in *The Canadian Journal of Philosophy* 41(2), June 2011

## **The Trouble with Being Sincere<sup>1</sup>**

Timothy Chan & Guy Kahane

---

<sup>1</sup> The authors are listed in alphabetical order; both contributed equally to the paper. We are very grateful to two anonymous referees for extremely useful comments. We have also benefited from comments by audiences in Oxford, Oslo and Geneva, where we presented earlier versions of this paper. Part of the work on this paper was supported by the Research Council of Norway and the Wellcome Trust (WT087208MF).

## The Trouble with Being Sincere

### Abstract

Which mental states of a speaker determine whether or not an assertion is sincere? Until recently the dominant answer has been that a speaker makes an insincere assertion if and only if one does not believe the proposition asserted. There are, however, persuasive counterexamples to this simple account. It has been proposed instead that an insincere assertion that  $p$  is one made by a speaker who (a) does not *express* his belief that  $p$ ; or (b) does not *believe* that he believes that  $p$ ; or (c) does not *assent* to  $p$ . We show that these alternative accounts also face counterexamples. We argue that, because of the disunity and opacity of the self, it is a mistake to identify insincerity with any privileged type of propositional attitude towards  $p$ . After diagnosing their failings, we sketch a new account according to which the operative state of mind is the audience-directed motivation of the speaker's behind the assertion. This motivation may, but need not, be expressed in the speaker's conscious intention; and sincerity may, but need not, require that one says what one mentally assents to. Thus defining sincerity in theory, as well as complying with the norm of sincerity in practice, both involve more trouble than might be expected.

Questions about sincerity play a central role in our lives. But what makes an assertion insincere? In this paper we argue that the answer to this question is not as straightforward as it has sometimes been taken to be. Until recently the dominant answer has been that a speaker makes an insincere assertion if and only if he does not believe the proposition asserted. There are, however, persuasive counterexamples to this simple account. It has been proposed instead that an insincere assertion that  $p$  is one made by a speaker who (a) does not *express* his belief that  $p$ ; or (b) does not *believe that he believes* that  $p$ ; (c) does not *assent* to  $p$ ; or (d) does not express any of these cognitive states. We show that these alternative accounts also face counterexamples. We argue that, because of the disunity and opacity of the self, it is a mistake to identify insincerity with any privileged type of cognitive attitude towards  $p$ . We sketch a different account according to which the operative state of mind is the audience-directed motivation of the speaker's behind the assertion. This motivation may, but need not, be expressed in the speaker's intention; and sincerity may, but need not, require that one says what one mentally assents to. Thus defining sincerity in theory, as well as complying with the norm of sincerity in practice, both involve more trouble than might have been expected.

### **I. The Basic Idea of Insincere Assertion**

Philosophers have been interested in sincerity of assertions for at least three reasons. Philosophers of language are interested in the state of mind which makes an assertion sincere, in order to delimit assertion as a type of speech act. Thus, for example, in Austin's (1975, 15) account of speech acts, insincerity is an infelicity that occurs when a linguistic convention designed for use by persons having certain thoughts or feelings is used when these thoughts or feelings are absent. Second,

epistemologists of testimony are interested in sincerity because it seems to be a necessary condition for the transmission of knowledge from speaker to hearer (Moran 2005, Owens 2006). Someone who asserts ‘*p*’ presents the asserted proposition that *p* as true. If the speaker did not even appear to accept its truth, the hearer would not normally take the assertion to be an epistemic reason in favour of *p* being true. And if the speaker merely *appeared* to take *p* as true, by making an insincere assertion, the assertion would normally not be a good reason for the hearer to accept its truth. Third, insincerity interests moral philosophers because in everyday life, as Ridge writes, “sincerity is important to us... because of the role it plays in our moral evaluations. ‘Be sincere’ is an important default moral norm.” Making an insincere assertion is at least *pro tanto* morally blameworthy because it “betrays a willingness to engage in deception” on the part of the asserter (Ridge 2006, 492).<sup>2</sup>

These considerations suggest the following schema:

**Basic Idea of Insincere Assertion.** An assertion is insincere iff the speaker’s *aim* in making the assertion is to present to the hearer as true what he *does not take* to be true.

The first element, the speaker’s *motivation*, encapsulates what is *pro tanto* morally objectionable about making an insincere assertion. An insincere assertion need not aim to make the hearer believe falsely — it might just aim to please or flatter the hearer.<sup>3</sup> The second element, regarding the speaker’s cognitive attitude, registers the

---

<sup>2</sup> See also Williams (2002, 97).

<sup>3</sup> Owens (2006, 113) and Ridge (2006, 494) present examples aimed at showing that insincerity needn’t involve an intention to deceive. But all these examples show is that the relevant motivation needn’t always aim to change the audience’s beliefs. The speaker may want to present ‘*p*’ as true for a range of reasons, which are not always to inform (or misinform) the hearer, or even to

fact that he presents the asserted content as true *despite* his not taking it to be true himself. Notice that an insincere asserter's *not* taking the assertion to be *true* need not mean that he takes it to be *false*—he might be agnostic, or even merely suspect it to be true.

In what follows we shall consider different ways of spelling out this basic schema. Past discussions have generally taken sincerity to be the primary notion, and are driven largely by linguistic and epistemological concerns. However, in what follows we shall focus on *insincerity*, which we understand, in line with everyday practice, to be primarily a moral notion, with a *pro tanto* tie to blameworthiness.<sup>4</sup> Even if insincerity can sometimes be justified, it has a clear moral significance, whereas sincerity as such is not praiseworthy. Now necessary and sufficient conditions for insincerity will also give us an account of sincerity *if* all assertions are either sincere or insincere. However, some of the accounts we shall consider adopt, either explicitly or implicitly, a plausible threefold categorization on which some assertions are neither sincere nor insincere. (Their sincerity status presumably has neither positive nor negative moral significance.)<sup>5</sup> We shall therefore assume in what

---

bring about the hearer's belief that the speaker believes that *p*. See Grice (1968, 123); Moran (2005, 328-330); Williams (2002, 71-2).

<sup>4</sup> For example, someone who is accused of making an insincere assertion would normally see the need either to justify himself by denying the claim, citing mitigating circumstances, or else apologize.

<sup>5</sup> See, for example, Ridge (2006, 495-501) and Williams (2002, 73), which imply that assertions which neither express the speaker's belief nor aim to misinform the hearer are neither sincere nor insincere. Arguably, assertions that aren't sincere but are clearly not blameworthy – e.g. saying what one doesn't believe in obedience to an order, in a play or in an ironic utterance (see Owens 2006, 113) – are better characterized as belonging to this third category, rather than insincere. When someone

follows only the truism that the sincerity of an assertion implies that it is *not insincere*.

Given that the negation of insincerity might not yet amount to a complete account of sincerity, and given that the notion of sincerity used in other theoretical contexts might not perfectly coincide with its moral sense,<sup>6</sup> the criticisms we shall offer of other views need to be qualified. What our criticisms *will* show, however, is that these views are insufficient as accounts of the moral notion of insincerity.

## II. The Belief Account

It's natural and common to assume that "take to be true" in the Basic Idea should be identified with "believe". Austin holds that, just as an insincere promise is "to say 'I promise' without intending, [to make an insincere assertion] is to say 'it is the case' without believing" (1975, 50). And Searle (1969, 65) offers the following influential view:

**(A) The Belief Account.** An assertion '*p*' is insincere iff the asserter does not believe that *p*.

Belief is the most natural candidate for the cognitive element of the Basic Idea. Ridge and others have argued, however, that it is subject to counterexamples.

First, it seems possible to assert sincerely what one does not believe, if one is mistaken about what one actually believes. Ridge gives a 'Freudian' example which

---

misunderstands an ironic utterance, we point out that the speaker didn't literally mean what he said, not that he was *insincere*.

<sup>6</sup> One reason to expect systematic divergence is that two primary foci of theoretical interest in sincerity are its role as a felicity condition partly constitutive of speech acts, and as a necessary condition for assertion to transmit knowledge. These are success conditions that make use of the positive notion of sincerity, whereas moral significance attaches primarily to *insincerity*.

involves someone having the repressed belief that his mother hates him.<sup>7</sup> This kind of example is exposed to the objections that it controversially relies on psychoanalysis, and that such literally pathological cases are a hazardous basis for generalization. But there is no shortage of more prosaic counterexamples to the Belief Account. As Peacocke notes,

Someone can make a judgment... but it not have the effect that judgments normally do – in particular, it may not result in a stored belief which has the proper influence on other judgments and on action. A combination of prejudice and self-deception, amongst many other possibilities, can produce this state of affairs. (1998, 90)

Consider the following everyday example of self-deception:<sup>8</sup>

**Example 1 The Chauvinistic Professor**

“Someone may judge that undergraduate degrees from countries other than her own are of an equal standard to her own, and excellent reasons may be operative in her assertions to that effect. All the same, it may be quite clear, in decisions she makes on hiring, or in making recommendations, that she does not really have this belief at all” (Peacocke 1998, 90).<sup>9</sup> Suppose the professor,

---

<sup>7</sup> Ridge (2006, 488-489).

<sup>8</sup> The points we make are compatible with most accounts of self-deception. All we need is the generally accepted assumption that what an agent consciously judges (or assents to) isn't always conclusive ground for attributing beliefs to her, and that such divergence may be motivated. (For an account of self-deception that relies on only these assumptions, see Audi 1982.)

<sup>9</sup> The professor in Peacocke's example believes that British universities are the best, and not merely, in Gendler's (2008) terms, “alieves” it. Alief is “automatic, arational” and generates action “without the mediation of ... desire” (557-8). By contrast, when the professor makes the hiring

a Briton, is asked in a newspaper interview about how the best American universities compare with the best British ones. In order to help oppose proposed cuts in British government higher education funding, she answers, “British universities are the best in the world”, even though she thinks this is not what an impartial and informed observer would say.

She does believe what she said, but her assertion is clearly insincere, given that it is motivated by an intention to deceive, and states what she judged to be false. Hence there are counterexamples to the Belief Account.

A natural response to this kind of example would be to observe that while the speaker both asserts and believes that  $p$ , the assertion and the belief are not appropriately related. The most promising candidate for the appropriate relationship required for sincerity is that the assertion must *express* the belief. In Owens’s influential account, “the expression of a belief is directly motivated by the belief expressed” and must be intentional (2006, 111), conditions that are not met in the professor’s case. This gives us:

**(B) The Belief Expression Account.** An assertion ‘ $p$ ’ is insincere iff the asserter does not express the belief that  $p$ , but merely presents herself as doing so. (Owens 2006, 113)

The Belief Expression Account is a considerable improvement on the Belief Account and embodies important insights, to which we shall return in the last section. However, while it handles the last counterexample, it does not give an adequate treatment of the following variant. Suppose now that, with no intention to hide her views from the reporter, the professor answers that British universities are *not* the best

---

decisions, for example, she acts consciously, deliberately, and in a way understandable by the usual rationalizing belief-desire explanation.

in the world, in accordance with her judgment (though not belief). She does not harbor any aim of deceiving the audience by her assertion, and we therefore would not criticize her assertion as insincere, contrary to what the Belief Expression Account says. The Belief Expression Account gives the wrong verdict that the professor's assertion is insincere because expressing a belief requires having the belief expressed. To elaborate, while the speaker's having the belief that *p*, as we have already seen, isn't sufficient for an assertion to express the belief that *p*, it *is*, on this account, a necessary condition. As Owens remarks, "you can't express beliefs you don't have" (2006, 109). Thus the Belief Expression Account agrees with the right-to-left conditional of the Belief Account: an assertion is insincere *if* the speaker doesn't believe what she asserts. As a result, both the Belief Account and the Belief Expression Account deliver the wrong verdict in the scenario in which our professor asserts what she judges to be the case. Owens responds that in such a case where "the speaker is not intending to misrepresent either the facts or the belief he holds", the assertion is "not a lie", but "not fully sincere either ... [since] lying is not the only form of insincerity" (2006, 126, Note 5). While not denying that it would be insincere for the professor to assert, contrary to her judgment, that British universities are the best, Owens claims that it would *still* be insincere, albeit to a lesser extent, for her to assert the contrary. But why, on the Belief Expression Account, should there be any difference (whether categorical or one of degree) in sincerity between the two assertions? After all, it's clear that *neither* assertion would express her belief; if, as Owens agrees, the norm of sincerity requires her to make one rather than the other, that must be because insincerity involves something not mentioned in the Belief Expression Account.

Another possible reply by both the Belief Account and the Belief Expression

Account would claim that the assertions involved are neither sincere nor insincere. The subjects in our examples are “of two minds” about the proposition asserted. Our professor, for example, believes that British universities are the best, but does not believe that she holds this belief, and is disposed to make the contrary judgment when she consciously assesses the evidence. Some may hold that, in such a mixed picture, it’s no longer determinate whether or not she has the belief.<sup>10</sup> It could thus be argued that it’s also indeterminate whether the assertion is sincere or insincere. Alternatively, one can put the examples above (and some of the ones we shall discuss below) under a third category of assertions that are neither sincere nor insincere. The problem with these proposals is that they falsely imply that it would make no difference, as far as sincerity is concerned, whether the professor affirmed or denied that British universities were the best. On the current proposal, it is indeterminate whether or not the professor believes British universities to be the best (or alternatively, she neither believed nor disbelieved it), and thus both assertions, coming out of her mouth, would have an equally indeterminate sincerity status (or both be neither sincere nor insincere). But the fact that her behavior outside the interview room shows her to believe it clearly doesn’t render her assertion, “British universities are the best” immune from moral criticism, when she deliberately sets out to give a false answer. So these proposals fail.

### **III. The Assent Account**

Instead of modifying the Belief Account, it might seem more promising to look beyond belief for the cognitive element of sincerity. The foregoing examples rely on the distinction between what Peacocke calls, in the passage quoted above, ‘stored

---

<sup>10</sup> Cf. Schwitzgebel (2001).

belief' and 'judgment', and the possibility that they can diverge. This distinction will be vital to our subsequent discussion, so let us look more closely at these notions. Peacocke understands belief in the broadly functionalist way which is standard in the contemporary literature, and which we shall assume here. Belief is a contentful mental state, typically but not always consciously accessible, whose content functions as a premise in the practical and theoretical reasoning that make sense of the actions and thought of the agent. It's generally accepted that attributions of beliefs must, in conjunction with attributions of other mental states such as desires, be able to explain an agent's actions. Thus, our grounds for saying that our professor really believes British degrees are the best would include evidence such as, "If she didn't believe that, why would she choose to hire A rather than B, who is obviously superior in all other respects?"

Whereas belief is a state of mind which persists over time without one's being aware of it throughout (if at all), judgment is "a conscious rational activity", "a mental action" (Peacocke 1998, 88). Peacocke's *judgment* is also known in the literature<sup>11</sup> as *mental assent*, or simply *assent*. One assents to '*p*' when one considers the question whether '*p*' is true, and gives an affirmative answer in thought; or when he affirms that *p* in his own mind with his interest in '*p*' being its actual truth value.

How are assent/judgment and belief related? Both belief and judgment, Peacocke observes (1998, 88), have the "fundamental goal" of aiming at truth, and, "to form a judgment is the fundamental way to form a belief (or to endorse it when it is being reassessed)". Many of our beliefs (notably perceptual beliefs) are formed without any process of conscious reflection, and thus without assent. However, when we *do* assent to '*p*', normally the belief that *p* is either thereby formed, or reinforced if

---

<sup>11</sup> Armstrong (1973); Mellor (1978); Price (1969); Shoemaker (1996).

it already exists. Conversely, when we believe that  $p$ , we are normally *disposed* to assent to ' $p$ ', in the sense that we *would* assent to it if our attention were directed towards answering the question whether it is true. This is no mystery, since if I judge that  $p$  is the case, it's only rational for me to draw inferences or act on the basis that  $p$  is true, which is to say to believe that  $p$ . Conversely, being inclined to assent to  $p$  is one central constituent (though only one) of the multifaceted dispositions that a believer of ' $p$ ' typically has. We are only imperfectly rational, however; so sometimes assent can fail to result in belief, and belief can persist in spite of dissent, as the examples above illustrate. Peacocke cites prejudice and self-deception as common causes of such failures of rationality; to the list we can add wishful thinking, weakness of will, psychological illness (as Ridge tells us) or simply carelessness. Thus we have cases of what Price calls 'unconscious beliefs', where "all or most of the other manifestations of a belief that  $p$  do occur in [the agent], but he does not assent to the proposition  $p$  when he entertains it and attends to it, and perhaps he even rejects it" (1969, 299-300). Conversely, someone who is disposed to assent to  $p$ , but lacks all or most of other such manifestations, doesn't believe that  $p$ , even though he may well mistakenly think he does. We call this 'empty assent'.<sup>12</sup>

Our counterexamples to the Belief Account all involve unconscious belief and/or empty assent. In such cases there are two incompatible candidates for what the speaker takes to be the case in the Basic Idea, and it's assent rather than belief which seems to fit the data better. Thus according to Shoemaker (1996, 78), for example,

---

<sup>12</sup> We assume that in order to understand assent, we need not appeal to the notion of sincerity (for example, sincerity of thought), at least not the same notion of sincerity as that which applies to assertions.

“[w]hen linguistic assent [i.e. assertion] is sincere, it involves mental assent.”

Correspondingly, we have:

**(C) The Assent Account.** An assertion ‘*p*’ is insincere iff the asserter is not disposed to assent to the proposition that *p*.<sup>13</sup>

Apart from enabling us to accommodate the counterexamples to the Belief Account, the Assent Account also has considerable independent appeal. What one assents to is what one consciously thinks is true, and (at least in ordinary circumstances) the speaker is conscious of what he is saying. Thus if the speaker asserts what he is not disposed to assent to, he is conscious of telling the audience what he doesn’t think is the case, which would typically amount to consciously intending to deceive the audience. On the other hand, when unconscious belief is involved, the speaker is by definition not aware of what he believes. Thus, even though what he says diverges from his belief, he is not aware that this is the case, and, it would seem, cannot be blamed for being insincere. Since intention, like assent, is accessible to the agent’s awareness by reflection, if we follow Williams (2002, 96) and others in spelling out the speaker’s aim in the Basic Idea in terms of the speaker’s *intention* to deceive the hearer, then it seems assent rather than belief should be the cognitive element when belief and assent diverge. Moreover, given that assent and belief normally converge, and that the ordinary language notion of ‘belief’ is often used to refer to the theoretical notion of assent as well as belief,<sup>14</sup> the Assent Account has a ready explanation of the plausibility of the Belief Account.

---

<sup>13</sup> The speaker needn’t actually undergo the mental act of assenting or dissenting to a proposition when he asserts it. What is relevant is that, if he *were* to consider the question whether *p*, he *would* not assent to it.

<sup>14</sup> See Cohen, 1992, §1.

In the next section we shall critically assess this account. Before we do so, however, we should say something about a different diagnosis of the foregoing examples that appeals to second-order belief rather than assent.

We have seen that Ridge thinks the Belief Account fails because of the examples we discussed. His alternative positive account is

**(D) The Second-Order Belief Account.** An assertion ‘*p*’ is insincere iff the asserter does not believe that he believes that *p*. (Ridge 2006. 501)

This proposal significantly overlaps with the Assent Account. Indeed, Mellor (1978, 97) takes the two to be identical when he writes, “Sincere assertion is saying what one assents to, that is, what one believes one believes, not just what one believes.” The two accounts deliver the same verdict over the vast majority of cases. For, as Evans (1982, 255) famously observed, we form second-order beliefs of the form ‘I believe that *p*’ by considering whether we assent to *p*. Certainly this is how we answer the question “Do I believe that *p*?” in ordinary circumstances; and to that extent, one assents to *p* if and only if one believes that one believes that *p*, and the Assent and the Second-order Belief Accounts coincide. However, there are contexts in which our second-order beliefs *don’t* seem to be formed in the way Evans describes. In these cases, assent and second-order belief diverge, and it’s the Assent Account which gives the correct verdict.

These examples also involve unconscious belief. Shoemaker, for example, discusses the following example, which is a variant of Ridge’s psychiatric case.

**Example 2 The Psychiatric Patient**

[S]uppose that a psychiatrist tells me that I have the repressed belief that I was adopted as an infant. In fact, the psychiatrist has confused me with another patient (he has been reading the wrong case history), and has no good grounds

for this belief attribution. But I accept it on his authority. It seems compatible with this that when I consider the proposition I am supposed to believe, that I was adopted, I find no evidence in its support, and am disposed to deny it.

(Shoemaker 1996, 89-90)

Now the patient assents to, “I was not adopted”, but believes, “I believe I was adopted”. He also thinks that his repressed belief that he was adopted is irrational, indeed pathological. The Assent Account predicts that he would be sincere if he were to assert, “I was not adopted”, and insincere if he were to assert the opposite; the Second-order Belief Account predicts that he would be sincere if he were to assert, “I was adopted”, and insincere if he were to assert the opposite. Clearly the former prediction is the correct one. For example, suppose the repressed belief at issue is “I was abused by my parents”, and the patient is asked in a court of law to give evidence about his parents’ alleged abuse. Given that he takes his (supposed) repressed belief that he was abused to be irrational and false, it would be insincere (indeed perjurious) for him to say he was abused. We shall therefore take the Assent Account rather than the Second-Order Belief Account as the one which best accommodates the examples we have discussed.

#### **IV. Against the Assent Account**

Unfortunately, the Assent Account delivers the wrong verdict on other examples. Let’s start with the case of Huckleberry Finn (Twain 1884), which has received much attention in moral philosophy since Bennett (1974). The story is set in early 19<sup>th</sup> century southern USA.

##### **Example 3 Huckleberry Finn and Jim**

Huck, a white boy in, endorses the conventional morality of his social milieu,

believing that slavery is just and that runaway slaves must be caught and punished. He has, however, befriended Jim, a slave on the run from his owner. Even though, not questioning the moral principles he had learned at school, he believes that it would be tantamount to theft to do so, Huck finds himself akratically helping Jim to escape.

As Arpaly & Schroeder (1999) argue, it seems not only that Huck acts in *conformity* with what is right, but that he deserves *moral praise* for his act. Or at least he deserves such praise if this is not some stray lapse, but an act driven by deep-set motivations, motivations genuinely responsive to the moral reasons Huck has to help a fellow human being, even if these are not consonant with his moral beliefs. As Arpaly and Schroeder write, even though “Huckleberry is not a good enough deliberator to reject his society’s deep convictions on the basis of his experience, ... the motive for his action is still a morally praiseworthy one” (1999, 164). They suggest that, all things considered,

an agent is more praiseworthy for a good action, or more blameworthy for a bad action, the more the morally relevant psychological factors underlying it are integrated within her overall personality. (1999, 172)

Consider now a version of the Huck example that involves *self-deception*. Suppose Huck has greater self-control. In this extension of Huck’s story, he isn’t the kind of person who often acts akratically, but the fellow feeling that inclines him to help Jim is still at work. Although he tells himself that he is really just keeping an eye on Jim, that his helpful acts are really a way of preventing Jim from disappearing, the way Huck behaves, and how he feels, indicates that this isn’t really what he believes. He is self-deceived about his actions: deep down he realises that he is helping Jim. Surely Huck still deserves moral praise, since there is still a core part of him that is

responsive to genuine moral reasons, and is a deeper expression of his self than what he consciously intends or assents to. Against this background, our story may quite plausibly conclude as follows:

**Example 3 (continued) Huck's Self-deceptive Assertion**

Huck knows that search parties are on their trail. Whenever Huck is troubled by the fact that he is doing something (he takes to be) wrong, he tells himself that the following day he will deliberately lead Jim into the pursuers' likely path. But each time he either forgets about the plan, or delays it under some flimsy pretext, while continuing on the path to safety. One night Huck tells himself that there is still time to carry out his scheme, but in fact (though he wouldn't admit it) he possesses enough information to know that they have left the pursuers behind. When Jim anxiously asks, "Do you think we have shaken off the search parties?" Huck answers, "Yes, we have indeed shaken them off," taking himself to be lying in order to keep Jim at his side until he turns him in. Jim feels reassured, and Huck finds himself strangely satisfied and relieved.

Just as in the original Huck example, the self-deceptive Huck deserves moral praise for his actions, and we do not see him as blameworthy simply because on a conscious but superficial level, his intention is to betray Jim. If we now focus in particular on his assertion to Jim that they are safe, and assess its sincerity status, we should reach the same verdict. Huck may not realise what he is doing, but his assertion really is motivated by a desire to inform Jim of an important piece of information. In a peculiar way, his self-deceptive assertion is driven by a desire to *reveal the truth* to Jim. Huck aims to inform Jim of what, deep down, he takes to be the case, even if he needs to go through elaborate mental acrobatics to do so.

To be sure, there is a sense in which we can say that Huck is trying to deceive Jim. But although Huck may take himself to be engaged in deception, his assertion isn't insincere. It is directly motivated by a desire to convey the truth. And this desire expresses an aspect of his self that is more central than his feeble conscious endorsement of the contrary aim — an aim which is a self-directed rationalization of the assertion rather than its causal engine.

If this interpretation is correct, then we must reject the Assent Account. Huck assents to the proposition “We have shaken off the search parties” yet asserts its negation. Nevertheless, his assertion seems to us clearly sincere. (In any case, in order to cast doubt on the Assent Account, it's enough if it is unclear whether our Huck is being *insincere*, a question that should be closed on that account.) For Huck also *believes* what he asserts, and his assertion is driven by the desire to inform Jim of the content of his belief. So Huck's assertion is *not* an instance of the Basic Idea of Insincere Assertion: Huck *does aim* to present to the hearer what he *takes* to be the case.

Now Huck's assertion is motivated by morally commendable motives, and it may be thought that we are reluctant to attribute insincerity to him, not because of any characteristic of the assertion itself, but because, overall, we approve of his actions to save Jim. But there are often enough situations when our judgment of the sincerity of an assertion and the overall moral status of the act diverge – we are familiar, for example, both with sinister sincerity and well-meaning lies. In such cases, asserting what one assents to may still be insincere, as our last example illustrates.

**Example 4 Susan the Faithful Daughter-in-law**

Susan's sea-faring husband Thomas was lost at sea years ago, and since the accident she has been taking care of his elderly and frail parents, who, despite

clear evidence to the contrary, has been clinging on to the hope that he is still alive. Susan has always been a terrible liar, and knows that Thomas's parents would be able to tell immediately if she merely pretended to think that he is alive. Motivated by her devotion to them, Susan focuses on the little evidence there is that Thomas might not be dead, and would mentally assent to his being alive. Nevertheless, her overall pattern of feeling and behaviour makes it clear that she is in fact reconciled to his being already dead.

If we are inclined to approve Susan's action, that is for the same reason that we would be inclined to approve if she wholeheartedly recognized the evidence, but consciously lied to the parents-in-law to protect their feelings. Insofar as we think that it's *prima facie* (though sometimes not all-things-considered) wrong to hide the truth from one's audience, it seems we have to give the same verdict to Susan's assertion, notwithstanding her assent to what she says.<sup>15</sup> For her assent is formed by a process of self-deception which is *itself* principally motivated by the desire to deceive others. Indeed, it's a familiar point that the easiest way to lie is to believe the lie yourself. When we consider such cases, it seems hardest to say that, just because the agent consciously assents to what she asserts, she must be sincere. For what explains why she thus assents, and why, consequently, she makes the corresponding assertion without any hesitation or discomfort, is precisely her unacknowledged aim *to deceive*.

In both our examples, what the speaker takes to be the case (in the sense relevant to sincerity) is what he believes to be the case, not what he assents to. But this does nothing to resurrect the Belief Account, or for that matter the Belief

---

<sup>15</sup> When you realize that someone has deceived you for your own good, you feel duped, even if you nevertheless approve of the act. Thomas's parents would have reason to feel the same, if they came to understand the full story behind Susan's actions.

Expression Account, since the counterexamples against these accounts still stand. Rather we now argue that, contrary to the four accounts we have considered, there is *no* privileged propositional attitude towards ‘*p*’ which is the final arbiter of what it is the speaker takes to be the case, the absence of which renders an assertion insincere.

## V. The Testimonial Motivation Account

As noted earlier, the accounts we have considered so far largely ignore the conative dimension of sincerity. It’s natural to assume that this conative dimension must be contributed by the speaker’s intention, as it is in Williams’s account. We saw earlier that everyday ‘belief’ talk is ambiguous between the dispositions respectively to assent to ‘*p*’ and to act on the premise that *p*, states that typically co-vary but can in some circumstances come apart. Our examples suggest that everyday talk about the speaker’s aim is also ambiguous, between his *intention* and what we may call his *motivation*. An agent’s intention, as already discussed, is akin to the disposition to assent in that it is accessible to awareness, even when it is not actually conscious. An agent’s intention in  $\emptyset$ -ing is primarily determined by what he is disposed to answer if he were to ask himself what his aim is in  $\emptyset$ -ing.<sup>16</sup> An agent’s intention typically, but not always, coincides with his motivation, just as the disposition to assent typically, but not always, coincides with his belief. And just as there are unconscious beliefs (i.e. beliefs contrary to assent), sometimes an agent’s motivation is unconscious, in the sense that the reason for which he acts is contrary to his (either conscious or

---

<sup>16</sup> See for example Anscombe (1957).

accessible-to-awareness) intention.<sup>17</sup> Correspondingly, in most – but *not all* – contexts we can determine the sincerity of an assertion by looking at the speaker’s intention. Hence, in the Basic Idea, neither ‘aim’ nor ‘take to be the case’ should be taken to refer to a privileged type of conative or cognitive state.<sup>18</sup>

This is a negative claim. What, then, should guide ascriptions of insincerity in cases where the self is divided in these ways? The acting agent’s intention is typically only one element among the complex of motives or conative states (including motivations that may or may not be in harmony with his intention) that explain his action. When an asserter’s intention is to misinform, this intention is typically driven by other motives—the agent’s reasons for wanting to misinform the hearer. These motives would feature in a complete explanation of why the agent made the assertion, but they would typically be compatible with the explanatory role played by the intention to deceive (or be truthful). For example, the motives behind the self-deceived professor’s statement may include beliefs about academic funding, her desire to impress the reporter, etc. These further attitudes operate either alongside or in conjunction with her conscious intention to inform or misinform. The cases of Huck and Susan, however, are different in that the broader motivations include a desire which is directly contrary to their intention to (respectively) deceive/tell the truth. Moreover, the unconscious motivation concerning whether to deceive or not, compared to the intention with the contrary content, plays a more fundamental causal role in motivating the assertion, in the following way. The intention here is *itself*

---

<sup>17</sup> Just as with ‘belief’, the everyday term ‘intention’ might also be used to refer to unconscious motivation. Our more narrowly defined usage allows us to distinguish importantly different phenomena—just as in the belief/assent case.

<sup>18</sup> Arpaly & Schroeder (1999, 171-2) similarly claim that moral praise and blame are not tied to any privileged mental state.

driven by the unconscious motivational state with contradictory content. Let us say that a motive to  $\emptyset$  is *dominated* by another motive *not* to  $\emptyset$  iff the latter not only exists, but causally explains the former by either bringing it about or sustaining it. As a result, a dominated motive to  $\emptyset$  would not have existed, had there not been the contradictory motive not to  $\emptyset$ . The motive to  $\emptyset$ , when it is a dominated one, is thus dependent on the motive not to  $\emptyset$ , but not conversely. We call an intention or motivation which is *not* dominated by a contradictory motive a *dominant* one. As a consequence of these definitions, when no contradictory motive not to  $\emptyset$  is involved in the explanation of an action, the motivation to  $\emptyset$  is dominant by default.

The foregoing analysis suggests that what constitutes the speaker's aim in the Basic Idea is the dominant motivation behind the assertion:

**(E) The Testimonial Motivation Account.** An assertion 'p' is insincere iff it is motivated by a *dominant desire* of the speaker's to present p to the hearer as true despite his (the speaker's) not taking p to be true.

A desire which is dominant needn't be deeply integrated into the agent's personality—after all, one may make an insincere assertion as a casual prank. This is a *testimonial* motivation account because the key conative question is whether the speaker desires to present to the hearer what the speaker takes to be the case. Whether what the speaker takes to be the case (i.e. the cognitive element in the Basic Idea) is believed or assented to depends on which cognitive state interacts appropriately with the dominant motivation. In Examples (1) and (2) and their variants, the speaker's intentions to tell the audience (what he takes to be) something true/false are dominant, in virtue of absence of contrary motivations. Their dispositions to assent are the relevant cognitive states because they combine with these (dominant) intentions to bring about the assertions. In Examples (3) and (4), on the other hand, unconscious

motivations dominate intentions, and it is beliefs that are relevant because they combine with the (dominant) unconscious motivations to bring about the assertions.<sup>19</sup>

When the agent's mind is not unified, the Basic Idea has no simple application. We need to decide which of two opposing motivational (or cognitive) states of the agent is relevant for a special kind of moral assessment. This point was already behind the move from the Belief to the Assent Account. In response to self-deception cases, it was natural to identify conscious assent (as well as intention) as determinative of insincerity. It might have seemed, at the time, that this was because, even if it's not always transparent to us what we believe, it should be at least transparent to us what we intend to assert, and whether we aim to deceive. Indeed, if these were not transparent to us, then it would no longer be under our direct control whether or not we are insincere. The foregoing discussions show that this is a mistaken interpretation of what is going on in these examples. What is doing the work is not transparency but dominance.

Once it's accepted that what we truly believe or want is not transparent to us, it's hard to see why we should assume that insincerity *is* self-transparent. Indeed, many ordinary cases of moral wrongdoing involve self-deception. Agents typically believe that what they are doing is right, or at least excusable, even when they commit the gravest crimes. Such agents have not simply made a cognitive error. They typically have all the relevant information at hand (indeed, they would reach the

---

<sup>19</sup> While in these examples we have intentions interacting with assent, and unconscious motivations interacting with unconscious belief to produce the assertions, these are not the only two possibilities. In examples having different structures, intention may interact with unconscious belief, and unconscious motivation may interact with assent to produce assertions. For example, I might judge that *p*, but, aiming to conceal this, blurt it out of some unconscious self-destructive motivation, despite not believing that *p*. We are grateful here to an anonymous reviewer.

correct moral conclusion about similar conduct by others). We don't hesitate to hold such agents responsible, indeed to subject them to serious punishment, although the wrongness of their acts isn't transparent to them. The agents in our examples are no different.<sup>20</sup>

There is one aspect of Huck's (and Susan's) self that takes what is asserted to be the case, and another that doesn't. So there is no simple application of the Basic Idea. We might therefore wish to say that the insincerity (or sincerity) of the assertion is *imperfect*. But as we have seen at the end of Section II, since the norm of sincerity does not cease to apply in such cases, there remains the question: On the basis of which of the speaker's propositional attitudes do we judge his assertion to be insincere? The Testimonial Motivation Account, we submit, provides a better answer than the alternatives.<sup>21</sup>

It is instructive before we conclude to compare the Testimonial Motivation Account with a modified version of the Belief Expression Account<sup>22</sup>:

**(F) The Generalized Expression Account.** An assertion '*p*' is insincere iff the asserter expresses neither his assent to '*p*' nor his belief (nor his second-order belief) that *p*.

---

<sup>20</sup> We don't deny that degree of consciousness can make a *difference* to accountability, but this isn't the same as seeing it as necessary for accountability. See Arpaly (2003, 159-162).

<sup>21</sup> We have so far deliberately focused on insincerity, and have been silent on lying. The notion of insincerity is clearly broader than that of lying, and seems to us to be the more fundamental one. But if insincerity simply is lying, then the accounts we have considered will also be accounts of lying. If this is so, then when Huck tells Jim, "We have shaken off the search parties", he only (mistakenly) thinks he is lying. One implication of our discussion, however, is that anyone who insists on tying lying to intention accessible to the speaker's awareness must accept that lying and insincerity can come apart—that Huck's assertion is a case of *sincere* lying.

<sup>22</sup> We are grateful to an anonymous reviewer for suggesting this alternative account.

Now nothing we have said is incompatible with the insight that sincerity involves the expression of one's cognitive states; indeed, insincerity on our account can arguably involve failure to express what one takes to be the case. But we are after an account of insincerity, and here we doubt that the notion of expression can do all the work, even on this more permissive account.

First, it is implausible that motivation plays no role in making an assertion insincere in the morally relevant sense. Indeed, it seems to us that lack of expression isn't sufficient for insincerity in the sense in question; without the relevant motivation, the resulting assertion might be non-sincere without being insincere. Possible examples might be cases where someone who is distracted or confused finds himself asserting offhand something he neither assents to nor believes. Such a speaker doesn't express either cognitive state since the appropriate relationship between assertion and either cognitive state is absent; yet (supposing the carelessness is itself unmotivated) he doesn't seem to be insincere (as opposed to non-sincere).<sup>23</sup>

Second, the key notion of expression needs to be modified and further developed in order to tackle the examples we have discussed. In the original version, as we have seen, expression is an *intentional* action. But our examples crucially involve unconscious motivation contrary to the agent's explicit intention – for example Huck takes himself to be telling a lie rather than expressing what he believes. This means that the notion of expression needs to be broadened to capture the

---

<sup>23</sup> Another possible counterexample is of someone who decides to assert that '*p*' just because he thinks this will flatter someone. He may then notice that, as it happens, he does, in fact, believe that *p*. But this speaker does not assert '*p*' *because* he believes it to be true, and would still assert it even if he didn't. Arguably, such assertions don't *express* belief or assent, but, whilst not admirable, they do not seem insincere.

appropriate kind of goal-directness when unconscious motivation is involved, so that someone may be said to express a state even while believing that he is not doing so.

Third, and more important, supposing that an adequate account of expression involving unconscious motivation is developed, there is further need to give a systematic account of the relationship between intention-based and unconscious motivation-based expressions, when the two conflict, as they do in the cases of Huck and Susan the daughter-in-law. In these cases we have *both* a motive to express the agent's belief/assent, *and* a contrary one to merely pretend to do so. Should we say, as seems natural, that Susan's assertion expresses her *assent*?<sup>24</sup> This would imply that expression of assent does not guarantee sincerity (since Susan's assertion is insincere), meaning that the Generalized Expression Account also fails to state *necessary* conditions for insincerity. And if we claim, with the original Belief Expression Account, that being motivated to pretend to express one's cognitive state suffices to make an assertion non-sincere, and indeed insincere, then we would be driven to the contradiction that Huck and Susan's assertions are both sincere and insincere. We have seen that, moreover, we cannot resolve such conflicts by always giving priority to either belief or assent. Thus the question remains how, in the cases we described, it's decided which of a person's different cognitive states is the one that is decisive in determining whether an assertion is insincere. The Generalized Expression Account's somewhat *ad hoc* appeal to a disjunction of cognitive states offers no help here. To the extent that this can only be decided by reference to the

---

<sup>24</sup> Note that, unlike the counterexample of confused offhand assertions, the Susan example — which we think is a variant of a common enough phenomenon — doesn't presuppose the distinction between insincerity and non-sincerity, or our background assumption that insincerity is primarily a moral notion.

dominant motivation, it seems that such motivation does play a constitutive role in the suggested account, leading to a view that coincides with our own proposal.

The self is not simple, and this complicates a range of ethical questions. When we try to ascribe moral praise or blame to someone, should we focus on their desires, second-order desires, will, evaluative beliefs, conscious aims, or some other mental state? We have argued that similar questions also arise for insincerity. One consequence of our discussion is that we might be insincere even when, from the inside, it appears to us that we are doing our best to present things just as we take them to be. In trying to be sincere we can't help but assert what we assent to, and in some cases this won't be enough. One consequence of the opacity of the self is that being sincere is no simple matter.<sup>25</sup>

---

<sup>25</sup> See Williams (2002) and Frankfurt (2005). We do not, however, agree with Frankfurt that the lesson from this is that "sincerity is also bullshit" (2005, 67).

## References

- Anscombe, E. 1957. *Intention*. Oxford: Basil Blackwell.
- Armstrong, D. M. 1973. *Belief, Truth and Knowledge*. New York: Cambridge University Press.
- Arpaly, N. 2003. *Unprincipled Virtue: An Inquiry into Moral Agency*. New York: Oxford University Press.
- Arpaly, N. and T. Schroeder. 1999. 'Praise, Blame and the Whole Self'. *Philosophical Studies* **93**: 161-188.
- Audi, R. 1982. 'Self-Deception, Action and Will'. *Erkenntnis* **18**: 133-158.
- Austin, J. L. 1975. *How to Do Things with Words*. New York: Oxford University Press.
- Bennett, J. 1974. 'The Conscience of Huckleberry Finn'. *Philosophy* **49**: 123-134.
- Cohen, L. J. 1992. *An Essay on Belief and Acceptance*. New York: Oxford University Press.
- Evans, G. 1982. *Varieties of Reference*. New York: Oxford University Press.
- Frankfurt, H. 2005. *On Bullshit*. Princeton: Princeton University Press.
- Gendler, T. S. 2008. 'Alief in Action (and Reaction)'. *Mind & Language* **23**:552-585.
- Grice, P. 1968. 'Utterer's Meaning, Sentence-Meaning, and Word-meaning'. Reprinted in his 1989 *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.
- Mellor, D. H. 1978. 'Conscious Belief'. *Proceedings of the Aristotelian Society* **78**: 87-101.
- Moran, R. 2005. 'Problems of Sincerity'. *Proceedings of the Aristotelian Society* **105**: 325-345.
- Owens, D. 2006. 'Testimony and Assertion'. *Philosophical Studies* **130**: 105-129.

- Peacocke, C. 1998. 'Conscious Attitudes, Attention and Self-Knowledge'. *Knowing Our Own Minds*. In C. Wright, B. C. Smith and C. MacDonald eds. New York: Oxford University Press: 63-98.
- Price, H. H. 1969. *Belief*. London: George Allen & Unwin.
- Ridge, M. 2006. 'Sincerity and Expressivism'. *Philosophical Studies* **131**: 487-510.
- Schwitzgebel, E. 2001. 'In-Between Believing'. *Philosophical Quarterly* **51**: 76-82.
- Searle, J. 1969. *Speech Acts*. New York: Cambridge University Press.
- Shoemaker, S. 1996. *The First-Person Perspective and Other Essays*. New York: Cambridge University Press.
- Twain, M. 1884. *Adventures of Huckleberry Finn*. London: Chatto & Windus.
- Williams, B. 2002. *Truth and Truthfulness*. Princeton: Princeton University Press.