

Deception and the Desires That Speak against It



## II—CHRISTOPH FEHIGE AND ULLA WESSELS

### DECEPTION AND THE DESIRES THAT SPEAK AGAINST IT

This article explores the role of desires in the ethics of deception. The argument concentrates on intrinsic desires not to have false beliefs and on the resulting role of false beliefs as building-blocks, not just causes, of harm. If there is a duty of beneficence at all (a duty to further welfare or a duty to avoid harm) and desire fulfilment is at least a component of welfare, there is often a direct wrongness in causing a false belief.

#### I

*A Conceptual Task and a Moral Task.* What is it for one person to deceive another? The quest for an answer that presents necessary and sufficient conditions has proved thorny. Matters that require attention include the epistemic status both of the deceiver and of the deceived, the means of deceiving and the distinction between committing and omitting, the shape and content of the deceiver's intentions, the directedness of the action at the deceived, and the individuation of the impact that the action has. The literature has become home to a considerable list of challenges that relate to those matters and of clauses that are designed to meet them.<sup>1</sup> There have been valuable insights, but one of them is that an adequate explication of deception, should there ever be one, is bound to be complicated.

The complications on the moral side are likely to follow suit, since for a moral assessment of deception every feature of the complex that an explication has identified as constituting deception needs to be assessed: which duties or values does the feature affect, and what is their relative weight? Moreover, each of those many questions has a direct and an indirect part. Each feature is an occasion to ask

---

<sup>1</sup> Fuller (1976), Mahon (2007), and Shiffrin (2019) present and discuss several of the issues and convey an impression of the size of the task. The task becomes even knottier when it is extended to self-deception, a topic that we are setting aside. We will examine the deceiving of others. Parts of the present article loom into epistemology; for bibliographical advice regarding those parts we are grateful to Jochen Briesen, Frank Hofmann, Eva Schmidt, Peter Schulte, and Verena Wagner.

whether the feature is bad or wrong ‘in itself’, but also whether it causes, typically or in a certain situation, something that is bad or wrong ‘in itself’.

The moral task is Herculean, and the only hope is to proceed step by step, treating one component at a time. The component we intend to look at lies at the core of deception—more precisely, in the epistemic impact that is a necessary condition of deception. If one person deceives another, that much is uncontroversial, she has an epistemic impact on that person. She causes one of the following things to happen in the deceived: a true belief weakens or vanishes or fails to come into existence, or a false belief strengthens, persists, or comes into existence. Among those, the persisting or coming into existence of the false belief is usually considered as the paradigm case, and it is the focus of this article too. We ask what is wrong with causing a person to have a false belief, regardless of specific ways of doing so, and we will assemble at least some elements of the answer.

## II

*Desire-Based Ethics, Molecularly, and Outwardness.* It makes sense to approach the question whether a particular item is right or wrong with tail-wind from a general theory of right and wrong. The theory we find most plausible sees the rightness and wrongness of actions rooted in desires. We shall sketch, but not stop to justify, how the theory takes up desires and then links their fulfilment to welfare, welfare to value, and value to obligation.<sup>2</sup>

*Desires as the origin.* People have intrinsic desires, which is to say: desires of something for its own sake, not of it as a means to something else. The characterization is not meant to exclude that in addition and, at the same time, a person has an extrinsic desire of the same thing; those would be two different desires. Henceforth in this article, whenever we write ‘desire’, we will mean ‘intrinsic desire’. Some such desires may be purely implicit in the sense that they are dispositions of, but not occurrences in, the desirer’s consciousness.

<sup>2</sup> For the step from desire fulfilment to welfare, see von Wright (1963, esp. §§5.9 and 5.11); Schaber (1997, §§1.3.5.a and 3.1.1.b); Carson (2000, ch. 3); and (for numerous further sources, too) Wessels (2011). For the subsequent steps from welfare to value and on to obligation, see Holtug (2003) and Bykvist (2010).

Provided that a person desires that  $p$ , the desire is fulfilled if and only if  $p$  is the case, and frustrated if and only if  $p$  is not the case.

*From desires to welfare.* A person's welfare is the fulfilment of her desires. Both 'welfare' and 'fulfilment' are quantitative notions, which take into account the number, duration, and strengths of desires. Cases in which it is bad for a person that she got something she desired are cases in which that thing conflicts with the fulfilment of other desires of hers.

*From welfare to value.* The value of a world is determined by its welfare profile: who in that world has how much welfare? We leave it open whether an adequate account of that value is utilitarian (only the sum of welfare counts), prioritarian (an increase in an individual's welfare adds more to the value if the individual is worse off to begin with), or egalitarian (an increase in the equality of the welfare levels adds to the value). It makes no difference to the value who is who (every two worlds that have the same anonymized welfare profile are equally good), and every additional crumb of individual welfare increases the value to some extent.

*From value to obligation.* Finally, obligation: you ought to bring about as much value as you can; it is obligatory and also right and permissible to do so, and it is wrong not to do so. You have various specific duties, too, specific pro-tanto duties and specific all-in duties, but they all hang together with the one grand duty; they are duties to  $\phi$  regarding those actions  $\phi$  that contribute value.

It will be handy to have a name for the entire doctrine. Since the term 'welfarist act-consequentialism with a desire-fulfilment theory of welfare' is a mouthful, we will use the term 'desire-based ethics' instead. Like every philosophical programme with grand ambitions, desire-based ethics faces its share of challenges. Questions can be asked, for example, about the metric for the strength of desires, the treatment of infinite populations, the relevant notion of 'bringing about' outcomes or value. In this article, we will assume that desire-based ethics is seaworthy.

Desire-based ethics is committed to molecularity and to outwardness, two features that are crucial for our inquiry. Both concern the states of affairs that ultimately bear value. Morality is molecular if at least some of those states are complex, and it is outward if at least some of them go beyond people's minds. If desire fulfilment is the currency of ethics, molecularity and outwardness ensue. Consider the states of affairs that according to desire-based ethics ultimately

bear negative value. They are of the form ‘Mary desires that  $p$ , and  $p$  is not the case’. Such a state is complex, and it can involve all kinds of extra-mental things since the desideratum, the state of affairs  $p$ , can involve all kinds of extra-mental things.

The significance of molecularity and outwardness is illustrated by well-known cases with regard to which desire-based ethics and hedonism reach different conclusions.<sup>3</sup> Hedonism will be understood here as the moral theory that has the same structure as desire-based ethics, but is centred around pleasure instead of desire fulfilment. One of the well-known cases of dissent is the pleasure machine, which lets its user ‘lead’ the rest of her ‘life’ in orgasmic fantasies that never get boring. Consider Mary, who has a strong desire to remain in touch with her real life, warts and all. We can connect her to the pleasure machine all the same, painlessly while she is sleeping, for the rest of her life. Nobody would find out that it happened against Mary’s will. Hedonists have no reason to refrain from that action; they are concerned with pleasure alone, which is maximized in the pleasure machine, and so the matter is settled. For those who endorse desire-based ethics, things look very different. On their view, Mary’s desire to remain in touch with her real family and friends and possessions and problems will provide a strong reason not to confine her to the machine, even if there will be less pleasure outside.

Another well-known case is the death-bed promise. Mary is dying and makes you promise that you will see to it that the novel she has written will be published, posthumously; the promise is a secret between you and her. Leaving aside the pleasure that the novel might give to others, hedonism knows of no reason for you to keep the promise. All that counts is the pleasure that Mary receives from believing that the novel will be published, and that pleasure will be

<sup>3</sup> Recent discussions of the first kind of case, connecting people to pleasure machines, include Hawkins (2016) and Lin (2016). The second kind of case, the death-bed promise, touches upon at least three different questions about the relation of desire fulfilment to welfare. Assuming that Mary desires that  $p$ , what difference does it make: (i) that Mary will never find out whether  $p$ ; (ii) that by the time  $p$  takes place Mary’s desire will no longer exist; (iii) that by the time  $p$  takes place Mary herself will no longer exist? Desire-based ethics as sketched in this section gives the same reply to all three questions: it makes no difference; the fulfilment is still a component of Mary’s welfare.

The literature on the three questions is formidable and intertwined. See, for instance, Carson (2000, §3.1, subsections 5 and 6); Feinberg (1984, §§2.3–2.7); Egonsson (1990, §§2.1.1 and 2.2); Mulgan (2007, ch. 4, subsection ‘Posthumous Harm’ of section ‘The Preference Theory’); Wessels (2011, §§2.2, 4.2, and 5.1); Dorsey (2013).

administered through the credible giving, not through the keeping, of the promise. Not so in desire-based ethics, where the focus will be on the fulfilling of her desire and thus on making real the content of the desire. And the content of her desires is that the novel be published—not just that she believes, or that she receives pleasure from believing, that it will be published.

The moral picture would alter if one showed some allegiance to desires but departed from desire-based ethics in the direction of an ‘experience requirement’ or a ‘belief requirement’ on welfare. One might claim, for example, that negative welfare does not have the shape ‘Mary desires that  $p$ , and  $p$  is not the case’, but the shape ‘Mary desires that  $p$ , and  $p$  is not the case, and Mary believes that  $p$  is not the case’. We mention claims of that kind because they have been considered by several authors and would present deceiving, breaking a death-bed promise, or connecting people to pleasure machines in a different light. However, we will not pursue the matter. An appeal to some such further requirement would go against the grain of desire-based ethics since it would weaken the commitment to desiredness as the feature that counts.<sup>4</sup> Such appeals fall outside desire-based ethics as it is characterized in this article.

Desire-based ethics claims that all that matters is people’s welfare without claiming that all that matters is how people feel. Welfare is more complex and reaches out into the world. Other doctrines respect those claims in other ways, but by our lights this is the most promising.

### III

*Desire-Based Ethics and False Beliefs.* With a normative framework in place, we zoom in on the paradigmatic feature of deception that we have set out to consider. In the form in which they are envisaged by desire-based ethics, molecularity and outwardness give us a special handle on the causing of false beliefs. The plan to limit the discussion by leaving aside specific *ways* of causing false beliefs remains in force.<sup>5</sup>

<sup>4</sup> Most authors who have looked at such extra requirements have rejected them; see the works by Carson, Mulgan, and Wessels referred to in the previous note.

<sup>5</sup> To illustrate the limitation: certain ways of causing false beliefs may lead to the erosion of trust and cooperation or may frustrate people’s desires to have the agent as a friend. In so

First we give a nod to a class of well-known considerations that are not part of the special handle. They concern the causal repercussions of the false belief. Our remarks on them will assume that there is a reasonable sense of ‘to cause’ that permits even followers of a Humean theory of action to say that a certain belief causes a certain action. For example, the sentence ‘Mary’s belief that the stores are still open causes her to go out’ can be true in roughly the sense that of the two separate entities, the belief and the going out, the first makes a decisive difference to the second—without the belief Mary would not go out. To say so is not to deny that some conative state is also involved and is a cause in the same sense.

Typically, a false belief has many causal repercussions, and some are morally relevant. Perhaps some effects will be good, some bad. Everything will depend on the specifics of the situation. In particular, how will the belief shape the believer’s actions, and how will those actions affect people? One class of blatant examples are crusades. They show that sometimes actions based on false beliefs will misfire badly and will take their toll on the agent and on others.<sup>6</sup>

Although causal observations can be found in the literature on the moral status of deception and are not our main topic, we include a plea here for one strand of them because it seems under-reported. The misfiring of actions due to false beliefs is not restricted to false beliefs about the effects of one’s actions on certain desiderata. The misfiring can also be due to false beliefs about the desiderata as such: what is it that I desire? The desires that count in matters of welfare are dispositions, and while I may have privileged epistemic access to parts of my own mind (for example, to the toothache I have while believing I have it), there is no reason to believe that I have such access to every disposition of my own mind. Moreover, the dispositions under consideration will have triggering conditions that

---

far as the drawbacks result from the said ways, they are not our topic. Many of them are covered in Sissela Bok’s tractate on lying (1978).

<sup>6</sup> William Alston (2005, p. 30) is particularly clear about the dangers of false beliefs: ‘[I]f we had beliefs . . . that were mostly false, we would constantly be led astray in our practical endeavors and would be unlikely to survive for long. . . . As for the higher life of the mind, it would become a chaos if we had to rely on mostly false beliefs. Our attempts . . . to create beauty . . . and to engage in fruitful and rewarding interactions with our fellows would be frustrated at every turn.’ Arguments for the negative instrumental value of false beliefs can also be obtained, with a modicum of transposing, from most arguments for the positive instrumental value of true beliefs. The arguments regarding true beliefs are provided by, for example, Hilary Kornblith (1993, §5), Michael Lynch (2004, p. 16), and Seana Shiffrin (2019, pp. 76 and 83).

are strong and special—for a desire that  $p$ , conditions of the type ‘if I fully, correctly, and vividly imagined  $p$  in a cool hour’. The stronger such conditions, the less privileged the access is that I have to my own desires. Somebody else may see more clearly than I do myself what *would* go on in my mind if the crucial conditions held.

The possibility that a person has false beliefs about her own desires is very much alive, and such beliefs, too, can cause actions—and can cause a sequence of actions and an entire way of life—that fall far short of best serving the person’s desires. The more the literature on deception focuses on examples of assorted personal interactions among a small number of individuals (two people talking about their love life, etc.), the more it risks overlooking some of the grossest deceiving of people about the content and relative strength of their desires. Such deceiving takes place on an industrial scale, and large sectors of the economy make it their business.

From one important kind of causal impact back to causal impact in general. Structurally, the considering of the causal impact of false beliefs is not germane to desire-based ethics. To be sure, desire-based ethics will process the causal impact on desire fulfilment, while—to continue with the prominent alternative from the previous section—hedonism will process the causal impact on pleasure, and that will make a difference. Still, the considerations are not only of the same kind, the causal kind, but will sometimes involve causal stories that are similar. Sometimes, for example, the way in which a false belief makes an action misfire, which then has effects on the fulfilment of people’s desires, resembles the story in which the same belief and misfiring has effects on people’s pleasure.

A significant difference lies elsewhere. Most people desire most of the time with respect to most states of affairs not to have a false belief regarding that state. They desire it in the ‘for its own sake’ sense in which we are using the term ‘desire’. In addition, they may well have instrumental attitudes against false beliefs, but that is neither here nor there. And most of the desires in question are purely implicit, but that, too, is neither here nor there. A purely implicit desire is a desire.

In view of the desires not to have a false belief, desire-based ethics will recognize a negative role played by false beliefs that is more direct. Mary desires not to have a false belief regarding some state of affairs, but she does. The false believing is *part* of a unit of the form ‘Mary desires that  $p$ , and  $p$  is not the case’—a unit that is a dose of

negative welfare and thus of negative value. There is not just badness that the false belief causes. There is a chunk of badness that the false belief co-constitutes.

If one passed over the badness co-constituted by false belief, one would get things wrong. Moreover, one would pass over a kind of badness that is epistemically more accessible than others. The accessibility claim is not that constituting is easier to detect than causing, but that in the badness we are looking at only one step is involved, rather than two: constituting only rather than causing and constituting. The earlier considerations were of the form ‘Mary desires that  $q$ , but her having some false belief will cause not- $q$ , and not- $q$  in turn will co-constitute the frustratedness of the desire, a frustratedness that is bad’. The pattern at hand is more direct in that the false belief relates to the frustratedness without a causal go-between. Detecting effects and detecting desires can both be tricky, but here is badness that we can ascertain at half the cost.

From the relation of false beliefs to desires *not to have false beliefs* we move on to the relation of false beliefs to desires *to have true beliefs*. Desires to have true beliefs are as the sand of the sea,<sup>7</sup> and their pervasiveness lends weight to the question whether our constitutive story carries over to them. The answer lies in the following connection: the presence of the false belief-whether- $p$  entails the absence of a true belief-whether- $p$ . It would be daring to deny that connection, not least because the most plausible characterizations of belief entail that it is impossible for one and the same person to believe that  $p$  and, at the same point of time, believe that not- $p$ . One consequence of the impossibility is that one cannot have, at the same time, a true belief-whether- $p$  and a false belief-whether- $p$ .

The connection suffices for the constituting. That the false belief is there entails that the desired true belief is not. And thus, given the desire to have a true belief-whether- $p$ , the false belief-whether- $p$  co-constitutes the frustratedness of the desire, a frustratedness that is bad. In sum, it makes little difference whether we look at false beliefs

<sup>7</sup> Perhaps the desire to know is an anthropological constant? William Alston writes: “All men by nature desire to know”, said Aristotle, and this dictum has been reaffirmed by many of his successors’ (2005, §2.1). Alston sees in our species curiosity and ‘a built-in drive to get the truth’ regarding quite a number of things. Similarly, Goldman (1999, §1.1) and Lynch (2004, pp. 15–18 and 120). We part company from Lynch where he goes beyond diagnosing that truth *is* often cared about for its own sake and claims that truth is ‘deeply normative’ and ‘worth caring about for its own sake’ (2004, p. 17, our emphasis; more in part 3 of Lynch’s book).

in relation to desires not to have them or in relation to desires to have true beliefs. The false beliefs play their part in badness either way.

In the literature on deception and lies, the badness that we are illuminating does not seem to receive the attention that it deserves. Seana Shiffrin, for instance, mentions the case of hearers, not addressed by the speaker, who form beliefs about issues that are ‘matters of curiosity to them’ but ‘are not germane to their welfare’ (2019, p. 79). Our point is that in some sense the curiosity-not-welfare constellation cannot occur. If you desire to have the information but have a false belief instead, that entire state of affairs constitutes a dent in your welfare, because welfare is desire fulfilment. Harm and badness are there. They may be outweighed by competing desires (like a speaker’s desire not to divulge a certain piece of information or desires that can only be fulfilled if the piece of information is not divulged—for such weighing, see also Shiffrin 2019, pp. 82 and 84), but they are there.

The path that we are pointing to can be understood as a third way in the morality of truth. Way one is to invoke truth-related duties of some kind or other (duties to keep promises, not to tell lies, not to deceive) that are not rooted in the value of welfare. To go that way, it seems to us, is to risk turning truth or truth-directedness into a fetish. Way two is to look at welfare as the only source of duties and to see false beliefs related to welfare only mediated through causation. To go that way, it seems to us, is to get things right on the first leg of the journey (which is about the connectedness of welfare and duty), but wrong on the second leg, which is about the specifics of welfare and thus of the duties relating to it. Our proposal, coming from desire-based ethics, is to set off on way two, but to take a different turn in the middle, away from the causal restriction.

#### IV

*Reflections on False Beliefs as Co-constituents of the Bad.* The wrongness of causing false beliefs, we have claimed, is anchored in axiology strongly and directly. The anchor is desire fulfilment. The approach invites various responses, two of which we will now address. One is a complaint that we have shown too much; the second, that we have shown too little.

© 2019 THE ARISTOTELIAN SOCIETY

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

*Aristotelian Society Supplementary Volume* XCIII

doi: 10.1093/arisup/akz012

‘Your train of thought applies to everything, not just to false beliefs’, one reaction might be. ‘You could have run the same argument about, say, childlessness. Since many people desire to have offspring, and the frustration of that desire is bad, those people’s childlessness, too, is co-constitutive of the bad.’ We reply that indeed it is and that we embrace the claim without hesitation. Since the items that do the work in our train of thought are quite general (a certain take on molecularity and a certain take on outwardness, both takes provided by desire-based ethics), so is the work that they do. When desires have the same structure, desiderata will have the same status. The ambition has been to show that, when it comes to badness, many cases of false beliefs play in a certain league, but not that most cases of childlessness play in a different league.

‘On the other hand, very little is gained by your exposition’, another complaint might run, ‘because the anchoring still requires desires not to have a false belief regarding this or that, and those are desires that people frequently have but may also fail to have. You have not put respect for truth on a robust footing.’ We acknowledge the central part of that analysis. Yes, the anchoring employs desires that are common but probably contingent and is to that extent probably contingent itself. We declare that necessity is not where the merit of the approach is supposed to lie. The merit is the directness, the non-causal immediacy, of the anchoring, not an unconditionality of the ground.

Moreover, we do not yearn for more ‘robustness’. The conditionality of the approach is a sensitivity to desires and so to welfare. To detach respect for truth from desires and welfare would be to jettison that sensitivity and, so it seems to us, to idolize truth. In that perspective, the unsupportedness of deontic claims that go further is a feature of desire-based ethics, not a bug.

All that being said, here is a word of caution. To deny that desires not to have false beliefs exist necessarily is to make a claim that is not trivial. Since sometimes conceptual truths come to light only after some digging, we should keep an open mind.

Regarding desires not to have false beliefs, the possibility of conceptual discoveries is illustrated by the following sketch of an argument. The sketch employs a crude dispositional conception of desire and a crude covariationist conception of belief. The crude dispositional conception of desire says that, with regard to a human being, a tendency to  $\phi$  is a desire to  $\phi$ . If a person has at a certain point of

time a tendency to eat chocolate, she has at that point of time a desire to eat chocolate—and so on. The crude covariationist account of belief says: a belief that  $p$  is the presence of some kind of indicator of  $p$  in the believer's brain or mind. The status of an item as indicator of  $p$  is constituted by the covarying of that item with  $p$ . For example, if a person tends to have a certain mental picture in her visual field if and only if there is a frog in front of her, the mental picture is an indicator of the presence of a frog.

With those thumbnail sketches in place (the theories they stand in for are considerably more complex), we turn our attention to Mary, who holds a false belief, and see an argument emerge that concerns her relation to that belief:

- (1) Mary falsely believes that there is a frog.
- (2) Mary pictures a frog although there is no frog, and Mary has a tendency to [picture a frog iff there is a frog].
- (3) Mary has a tendency to [picture a frog iff there is a frog].
- (4) Mary has a tendency [not to picture a frog if there is no frog].
- (5) Mary has a tendency [not to falsely believe that there is a frog].
- (6) Mary has a desire [not to falsely believe that there is a frog].

Claim (1) is an assumption. Claim (2) follows from claim (1) and the crude form of covariationism about belief that we mentioned. We are sticking to the example of the mental picture as the indicator in order to keep things simple, but the argument would also go through, in a slightly more cumbersome form, if we used an existential quantifier instead: 'There is an item  $i$  in Mary's mind or brain of which the following holds true: . . .'. Claim (3) is a logical weakening of claim (2). Claim (4) follows from claim (3) provided that tendencies are closed under entailment. The same closure principle is at work in the step from claim (4) to claim (5); the relevant entailment (not picturing a frog entails not believing that there is a frog) is provided by the covariationist conception of belief. Claim (6) follows from claim (5) and the crude form of dispositionalism about desires that we mentioned.

To those familiar with the literature on mental representation, claim (5) of the argument won't come as a surprise. The claim that beliefs tend to be true or that most beliefs are true is not unusual.

Semantic externalists, in particular, find themselves drawn to the view that the mechanics constitutive of something being a content ensures that content comes with truth a lot of the time.<sup>8</sup>

The argument from claim (1) to claim (6) would show that, if Mary has the false belief that there is a frog in front of her, she desires not to have the false belief that there is a frog in front of her. Since Mary and the state of affairs were arbitrary (as was the example of an indicator), the argument generalizes. It would show: if a person has a false belief, she desires not to have that false belief. There is a frustrated desire wherever there is a false belief.

It is not our ambition to elaborate or defend the argument. Maybe the crude conceptions of desire or of belief are on the wrong track, or maybe they are on the right track, but the full-grown defensible versions of them have features that make all the difference. Maybe the closure principle for tendencies is misguided. The suspicion that at least one roadblock will turn up is not alien to us. We have no quarrel with that possibility because our point in this article is the sensitivity to desires and not the necessity of desires. The purpose of presenting an argument for necessity is methodological. There is at least a candidate for the role of a sound argument establishing that, necessarily, false beliefs come with frustrated desires. The candidacy suffices to show that one needs to proceed with care when denying that necessity itself or other necessities nearby. The careful procedure is to X-ray belief and to X-ray desire and to see what results come up.

## V

*Questions regarding the Bifurcated Approach to Deception.* How do the thoughts we have put forward relate to existing work on the concept and the morality of deception? We cannot do justice to the debate here, but can at least consider one contribution rather than

---

<sup>8</sup> Robert Stalnaker (1993, esp. pp. 299–305) explains the general outlook and the connection to belief. Given the ‘information-theoretic account of intentional content’, he writes, it ‘is reasonable to assume that representational states are *normally* correct—that they are states that *tend* to represent things as they are’ (p. 302). And, about beliefs in particular, ‘Certain of my internal states are beliefs about water because they are normally sensitive to facts about water. Normally, I would not be in the state I describe as believing that there is water in the vodka bottle if there weren’t water in the vodka bottle’ (p. 304). Bernard Williams (1970, esp. p. 108) expresses similar thoughts.

none. Seana Shiffrin's article 'Learning about Deception from Lawyers' is particularly rich and deserving.

Shiffrin is looking for a characterization of deception that respects at least four claims (2019, §§I and II), which for present purposes we have named and simplified.

*Unintentionality.* Sometimes there is deception even though the speaker does not intend to cause the false belief.

*Unreasonableness.* Sometimes there is deception even though the hearer's drawing the false inference from the speaker's behaviour is unreasonable.

*Limitedness.* Sometimes there is no deception even though a hearer predictably draws the false inference from a speaker's utterance.

*Permissibility.* Some forms of deception are permissible.

We will refer to the cases mentioned in the first claim as cases of 'unintentional deception'. The terminology is a little risky,<sup>9</sup> but so convenient that we won't resist.

For attempts to explicate deception, the combination of claims is a hard nut. On the one hand, there are many reasons not to simply equate deceiving with the causing of a false belief. If, for example, you introduce Peter to Mary and ten months later they have a child who five years later believes in the existence of elves, you have caused a false belief, but have not deceived. Leaving aside the example, the equating at issue is also blocked by *Limitedness*. We thus need to look for some more specific feature of the interaction. On the other hand, *Unintentionality* and *Unreasonableness* make it hard to see where that feature could reside. While *Unintentionality* makes

<sup>9</sup> The terminology competes with an 'as-such terminology', which might be seen as logically more proper. Let us assume, following Shiffrin, that an adequate characterization of deception is complex. In the as-such terminology the term 'intentional deception' would be used for all cases of deception in which the deception as such, which is the entire complex, is intended—that is, for all cases of the kind 'acts in that complex way, and intends to act *in that complex way*'.

For example, if the as-such terminology is used in connection with a disjunctive characterization of deception (like Shiffrin's) that has intending to cause the false belief as the first disjunct but no intentions in the other, then to call a deception intentional is to say that the deceiver intended to make the disjunction true. Two major differences between the terminologies would then be that, according to the as-such terminology, (i) intentional deceiving can take place without the intention from the first disjunct, but (ii) cannot take place without a nested intending in the following sense: an intending to  $\phi$  for some complex  $\phi$  in which to intend to cause the false belief plays some logical role.

We have explained the caveat for the positive term, 'intentional deception', since Shiffrin uses that term but not the negative one, 'unintentional deception'. The caveat applies to both terms alike.

it hard to find the distinctive feature on the speaker's side (we can't appeal to her intention to cause a false belief), *Unreasonableness* makes it hard to find it on the hearer's side (we can't appeal to her reasonableness in adopting the false belief).

Does it help to moralize things? Is 'deceiving' (like 'being cowardly') a thick ethical concept, with a moral dimension built in? The wrongness itself might then be the missing ingredient—the feature that turns the causing of a false belief into deceiving. That option, however, is blocked by *Permissibility*.

We are in a quandary, and Shiffrin has an ingenious proposal to get us out. She goes disjunctive, presenting 'the bifurcated approach':

One deceives when [(i)] one leads another by intention to form or reinforce a false belief *or* ... [(ii)] one leads another, in a relevant way, to form or reinforce a false belief by failing to fulfil a duty to take due care toward the other's mental contents. (Shiffrin 2019, §II)

The connection to the four claims is roughly as follows. Disjunct (i) effects compliance with *Permissibility* by failing to mention a breach of duty, while disjunct (ii) effects compliance with *Unintentionality* by failing to mention intentions. The duty of due care in (ii) is expected to have the right shape not only to leave all that complying intact but also to effect compliance with the remaining two claims, *Unreasonableness* and *Limitedness*. The duty makes the characterization of deception steer a middle course, with a chance of preventing the concept from applying in either too many or too few cases of errors on the receiving side. In particular, the duty sorts the cases in which your hearer predictably draws an unreasonable inference from your utterance and thereby acquires a false belief that you didn't intend to cause; in some of those cases your utterance was a violation of your duty to take due care (deception), and in others it wasn't (no deception).

There is much in Shiffrin's treatise that we endorse, not least her wake-up call regarding unintentional deception. We endorse her four claims, which she supports with an impressive array of examples and general considerations, and also her analyses of the difficulties that the quartet poses for various kinds of theories. What about her own proposal, though, the bifurcated approach? We find it original and forceful, but need to know more before we subscribe.

Our first question about the bifurcated approach is: what happened to unintentional deceiving that is not wrong? The bifurcated approach excludes that such a thing exists (Shiffrin 2019, p. 79), but *why* would one want to exclude that? A tendency to do so would be more understandable if it emanated either from the more general view that unintentional deceiving doesn't exist or from the more general view that permissible deceiving doesn't exist—but those are views that Shiffrin emphatically denies, with the denials being two of her four benchmark claims. So the account has a major structural feature that awaits explaining. If deception can be unintentional and can be permissible, why can it not be unintentional and permissible?

Shiffrin makes it clear that she is particularly interested in 'negligent deception' (2019, introduction and §1), where the negligence comes with culpability and liability. The topic is fascinating and deserves to be put on our radar. However, an honourable interest in studying one kind of deception—deception that is unintentional and wrong—does not justify the excluding of one of the other kinds from a general characterization of deception.

Examples from the excluded category do not seem hard to come by once the existence of unintentional deception is acknowledged, as it is by Shiffrin. We can pick any acknowledged example of unintentional deception and modify or supplement it so that nobody's welfare is adversely affected. The deceived person doesn't mind having false beliefs concerning the domain at issue, would not profit in any other respect from having true versions of them, doesn't mind the way the belief has been brought about—and so forth.<sup>10</sup> There is no

<sup>10</sup> Let's apply the construction manual to one of Shiffrin's examples (2019, pp. 73–4). An advertisement for the drug Anacin-3 contained the sentence 'Hospitals recommend acetaminophen, the aspirin-free pain reliever in Anacin-3, more than any other pain reliever'. We follow Shiffrin in assuming, for the sake of the argument, that the author of the statement did not intend to deceive. Unsurprisingly, the advertisement 'had the tendency to lead audience members to draw the false conclusion that Anacin-3 was the *brand* prescribed most by hospitals'. Here we have a case, says Shiffrin, 'of negligent, wrongful deception' (p. 73; similarly, pp. 77–8).

We can now strip the story of all parts that have moral significance. Say that the advertisement was authored and paid for by the discoverer of the beneficial effects of acetaminophen, who enjoys spreading the word. The hearers have no curiosity in any of these matters. Their false belief about Anacin-3 will never inform any decision. And so forth. It is hard to resist the following conditional judgement, the antecedent of which Shiffrin endorses: *if* deception can be permissible at all *and* the utterance is an unintentional impermissible deception in the original case, then it is an unintentional permissible deception in the modified case.

harm done, and thus no wrongness. Does Shiffrin believe that the ‘thus’ does not hold and that all other moral roads to relevant examples are also closed? The answer to that question bears a lot of weight and is worth giving and supporting.

A second question concerns the moral status of an agent’s deliberation. Shiffrin presents the example of Bette, who has breast cancer and for a while deceives close friends about her state (2019, p. 80). Bette does so thoughtfully in one case and thoughtlessly in another. If we understand Shiffrin correctly, according to her view there could be a pair of such cases that has the following properties. On the factual level, the two cases differ only with respect to the speaker’s thoughts. The utterance as such is the same in both cases, and the impact of the speaker’s mind and actions on people’s beliefs and welfare is also the same. The justification for the deception that would come to the speaker’s mind if she considered the question of justification would also be the same, and equally valid, in both cases. However, the speaker has in mind the justification in the first case, but not so in the second, in which she is simply not bothered about the justifiability. And although that is the only difference, the deceiving is right in the first case, but wrong in the second.<sup>11</sup>

In asserting the moral difference, Shiffrin goes beyond the familiar claim that thoughtlessness need not annul wrongness (a behaviour that is wrong when performed thoughtfully can remain so when performed thoughtlessly) on to the more audacious claim that thoughtlessness can *create* wrongness: a behaviour that is permissible when performed thoughtfully can become wrong just by being performed thoughtlessly. In Shiffrin’s view there are thoughts or maxims or motives that have some right-making or wrong-making force of their own, regardless of their impact on anybody’s welfare. That is a view we find hard to share. And the claim that the thoughtless person fails to take ‘due care’ doesn’t make it easier for us, because we don’t see where care that would have no impact gets its due from.

Is our understanding of Shiffrin’s position incorrect, and does her moral assessment of thoughtlessness stem from worries about the resulting welfare profile after all? If so, in what way? Perhaps the instance of thoughtlessness *is* bad just because the instance *could* have bad consequences, although it doesn’t? To us, such a ‘because’ from

<sup>11</sup> In some respect, the pair of cases mirrors the pair from our previous footnote. In the previous pair we kept the unintentionality constant and varied other features, whereas in this pair we are keeping other features constant and are varying the speaker’s thoughts.

‘could’ to ‘is’ would look precarious. We agree that carrots or sticks may be called for, incentives for a thoughtless person to become thoughtful, but if they were called for in Bette’s second case, they would be so, it seems, due to the risk of the badness or wrongness of future instances of thoughtlessness and not to any badness or wrongness of the instance at hand.

Our final question is related but more general. It concerns the structure of the duty that looms large in the bifurcated account, the ‘duty to take due care toward the other’s mental contents’. Shiffrin warns us that the duty is subject to questions that are ‘unwieldy’ and to ‘contextual complexities’; the moral territory lacks ‘tidiness’ and is ‘dauntingly rugged’ to an extent that may ‘evoke despair’ (2019, pp. 84 and 88–9). Witness Shiffrin’s own discussion of the duty, which touches on the deontics of acts versus omissions, autonomy, cooperation, demandingness, dignity, distributive justice, division of labour, the flourishing of conversations and people, privacy, roles, self-discovery, the speaker’s deliberative process, and special relationships.

We agree that the duty to take due care towards the other’s beliefs is complicated in the way in which applied ethics almost always is. However, we wonder whether Shiffrin goes further and espouses a particular kind of ruggedness that is more controversial: ruggedness due to an ultimate heterogeneity of the deontic world. Are issues that pertain to the forming of people’s minds—autonomy, friendship, omissions, privacy, and numerous further issues—governed by a motley crew of free-standing duties, which are not all moored to one general consideration? Perhaps some of the despair that Shiffrin mentions can be avoided if a more systematic picture is adopted, one that is more unified and deductive. The outlook that we sketched in §II has every particular duty originate in the value of desire fulfilment. We are under no illusion about the smoothness of the path from such an outlook to judgements regarding specific actions in specific situations. Quantification, verification, and computation are waiting to be performed. However, the outlook provides a common currency. It tells us what to look for when we seek to establish pro-tanto duties and, provided one of the possible formulae for aggregation is plugged in, how to weigh the pro-tanto duties in order to establish all-in duties. Is the bifurcated approach open to such benefits? Does the approach permit us to see its mighty moral

component, the duty in clause (ii), in the light of an ultimate moral homogeneity?

## VI

*Conclusion.* So much for an attempt to make some headway. The attempt has been to sketch a general framework, desire-based ethics, and to develop from it one thought in particular: often people's desires not to have false beliefs make it wrong to induce false beliefs. When a false belief about a matter is added to a desire not to have a false belief about the matter, the belief completes a desire frustration and thus a piece of negative welfare. The false belief completes a harm. You ought not to induce the false belief because you ought not to harm the believer. Those ought-judgements hold other things being equal and, since they have weight, often all things considered, too.

We have developed the central thought from one far-reaching doctrine, according to which desires hold two monopolies. They hold the monopoly regarding welfare, and, since welfare in turn holds the monopoly regarding moral value and moral obligation, they hold the monopoly in morality. However, the central thought has a wider range and also reaches outlooks that acknowledge neither of the two monopolies. Consider the claim that desire fulfilment is at least a component of welfare, possibly among other components, and that there is some duty, possibly among other duties, to further people's welfare. Suppose that either that pair of claims is in place or the pair of their negative correlates: desire frustration is one kind of harm, and there is a duty not to harm. Those pairs of claims are not exotic, and each pair suffices. If that ground is there, the central thought has traction.

The thought is both simple and potent, but also in danger of not getting its due in the ethics of lies and deceptions. There is a tendency to look elsewhere. In matters of truth, some see little occasion to consult welfare-directed obligations because they believe in truth-directed obligations that are independent of welfare. And some of those who consult welfare-directed obligations see little occasion to consider desires concerning truth or falsehood because they consider various other desires instead, desires that apply indirectly in that

their frustratedness would be caused, not constituted, by false beliefs or by certain ways of bringing about false beliefs. The desires directed at truth and the desires directed against falsehood should be part of the picture.

*Department of Philosophy  
Universität des Saarlandes  
Campus, Building A2.3  
66123 Saarbrücken  
Germany  
christoph@fehige.info  
ulla@uwessels.de*

## REFERENCES

- Alston, William P. 2005: *Beyond 'Justification': Dimensions of Epistemic Evaluation*. Ithaca, NY: Cornell University Press.
- Bok, Sissela 1978: *Lying: Moral Choice in Public and Private Life*. New York: Pantheon Books.
- Bykvist, Krister 2010: *Utilitarianism: A Guide for the Perplexed*. London: Continuum International.
- Carson, Thomas L. 2000: *Value and the Good Life*. Notre Dame, IN: University of Notre Dame Press.
- Dorsey, Dale 2013: 'Desire-Satisfaction and Welfare as Temporal'. *Ethical Theory and Moral Practice*, 16(1), pp. 151–71.
- Egonsson, Dan 1990: *Interests, Utilitarianism and Moral Standing*. Lund: Lund University Press.
- Feinberg, Joel 1984: *Harm to Others*. Oxford: Oxford University Press.
- Fuller, Gary 1976: 'Other-Deception'. *Southwestern Journal of Philosophy*, 7(1), pp. 21–31.
- Goldman, Alvin I. 1999: *Knowledge in a Social World*. Oxford: Oxford University Press.
- Hawkins, Jennifer 2016: 'The Experience Machine and the Experience Requirement'. In Guy Fletcher (ed.), *The Routledge Handbook of Philosophy of Well-Being*, pp. 355–65. Abingdon: Routledge.
- Holtug, Nils 2003: 'Welfarism: The Very Idea'. *Utilitas*, 15(02), pp. 151–74.
- Kornblith, Hilary 1993: 'Epistemic Normativity'. *Synthese*, 94(3), pp. 357–76.
- Lin, Eden 2016: 'The Subjectivist List Theory of Well-Being'. *Australasian Journal of Philosophy*, 94(1), pp. 99–114.

- Lynch, Michael P. 2004: *True to Life: Why Truth Matters*. Cambridge, MA: MIT Press.
- Mahon, James Edwin 2007: 'A Definition of Deceiving'. *International Journal of Applied Philosophy*, 21(2), pp. 181–94.
- Mulgan, Tim 2007: *Understanding Utilitarianism*. Stocksfield: Acumen Publishing.
- Schaber, Peter 1997: *Moralischer Realismus*. Freiburg and Munich: Verlag Karl Alber.
- Shiffrin, Seana Valentine 2019: 'Learning about Deception from Lawyers'. *Proceedings of the Aristotelian Society Supplementary Volume* 93, pp. 69–90.
- Stalnaker, Robert 1993: 'Twin Earth Revisited'. *Proceedings of the Aristotelian Society*, 93(1), pp. 297–311.
- Wessels, Ulla 2011: *Das Gute*. Frankfurt am Main: Vittorio Klostermann.
- Williams, Bernard 1970: 'Deciding to Believe'. In Howard E. Kiefer and Milton K. Munitz (eds.), *Language, Belief, and Metaphysics*, pp. 95–111. Albany, NY: State University of New York Press.
- von Wright, Georg Henrik 1963: *The Varieties of Goodness*. London: Routledge and Kegan Paul.