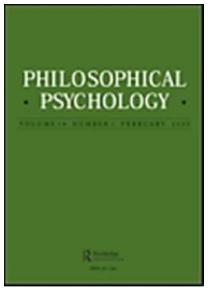

How is Rationalization Possible? A Model of Rationalization as Performative Pretense

Journal:	*Philosophical Psychology*
Manuscript ID:	CPHP-2013-0204
Manuscript Type:	Original Paper
Keywords:	rationalization, belief, pretense

How is Rationalization Possible? A Model of Rationalization as Performative Pretense

Abstract:

Rationalization in the sense of self-justification is a phenomenon very familiar to us all. *It's not cheating because everyone else is doing it too. I didn't report the abuse because it wasn't my place. I understated my income this year because I paid too much in tax last year. I'm only a social smoker, so I won't get cancer.* The mental mechanisms subserving rationalization have been studied closely by psychologists. However, when viewed against the backdrop of philosophical accounts of the regulative role of truth in doxastic deliberation (deliberation about what to believe), rationalization can look very puzzling. Almost all contemporary philosophers endorse a version of the thesis of deliberative exclusivity: a thinker cannot in full consciousness decide whether to believe that *p* in a way that issues directly in forming a belief by adducing anything other than considerations that she regards as relevant to the truth of *p*. But, as I argue, rationalization involves the weighing of considerations that the thinker knows very well are truth-irrelevant or inconclusive. This paper reconciles rationalization with deliberative exclusivity by modeling rationalization as a kind of performative pretense.

How is Rationalization Possible? A Model of Rationalization as Performative Pretense

To some the question, "How is rationalization (in the sense of biased self-justification) possible?" may sound otiose considering how familiar and pervasive the phenomenon is. To have an intuitive sense of the susceptibility of deliberation to the influence of non-truth-relevant considerations one need not be acquainted with the voluminous psychology literature on the cognitive mechanisms underlying motivated reasoning. *It's not cheating because everyone else is doing it too. I didn't report the abuse because it wasn't my place. I understated my income this year because I paid too much in tax last year. I'm only a social smoker, so I won't get cancer.* This kind of self-serving "logic" is such commonplace that it quickly become banal.

However, when viewed against the backdrop of philosophical accounts of the regulative role of truth in doxastic deliberation (deliberation about what to believe), this kind of reasoning demands analysis and explanation. In the debate about the "aim of belief" a rare point of agreement between normativists[1] (who hold that it is a conceptually constitutive normative feature of beliefs that they ought to be true), teleogists[2] (who hold that belief aims at truth in the psychological sense that beliefs are intended by agents or regulated by sub-personal mechanisms to be true), and skeptics[3] (who hold that various formulations of the aim thesis are false or platitudinous) is that from the perspective of first personal doxastic deliberation, only considerations that appear relevant to the truth of the proposition being considered can have any influence on the deliberative outcome. Indeed, the many of participants in the contemporary debate take it as an important desideratum that their theories account for this aspect of the phenomenology of doxastic deliberation. When I deliberate about whether to believe that *p*, it makes no difference what I feel I morally ought to believe, nor what practical aims I might have, nor what it would be most pleasant to believe. From this perspective, whether to believe a proposition is just a matter of whether the proposition is true. I will call this feature of doxastic deliberation "exclusivity"[4].

In light of the near-consensus[5] regarding the exclusive relevance of truth in deliberation over belief, it might seem as if doxastic deliberation promises to inoculate us from the influence of wishful thinking so common in non-ratiocinative formation of belief. If it is the case that only considerations that appear relevant to the truth of a proposition can play a role in our deliberation about whether to believe it, then considerations formulated with an eye to expediency ought to be excluded. Nevertheless, it is manifest that this is not the case. Far from providing inoculation to the influence of wishful thinking, protracted deliberation provides occasion for elaborate and highly strategic forms of it.

[1] See Boghossian (2003), Engel (2013), Fassio (2011), Gibbard (2005), Shah (2003), Shah and Velleman (2005), Wedwood (2002).
[2] See Velleman (2000), McHugh (2011) and Steglish-Petersen (2006, 2009, 2011).
[3] See Owens (2000, 2003), Glüer and Wikforss (2009), Hazlett (2013)
[4] I follow Steglich-Petersen (2009) in adopting this term. Exclusivity is endorsed by Bennett (1991), Walker (1996, 2001), Owens (2000, 2003), Kelly (2002), Shah (2003), Hieronymi (2008), and Steglich-Petersen (2009) among others.
[5] McHugh (forthcoming) dissents.

Of course, no one explicitly denies that deliberative belief-formation can be influenced by non-epistemic factors. To do sound would fly in the face of common sense. If, for example, it is comforting for a person to believe that he has earned his position by merit rather than family connections, the pleasantness of the belief may well play a causal role in bringing about his deliberatively-formed endorsement of it. Conversely, if it is painful for a person to believe that others think that he is undeserving of his office, then this "pain" might explain his slowness to revise the belief that he deserves what he has in the face of countervailing evidence.

But exclusivity implies that this kind of causal influence will not go via the content of the considerations adduced in deliberation when those considerations are known to be irrelevant to the truth. The truth-irrelevant deliberative considerations will not move the thinker in the guise of reasons. Although a thinker may form an intention to cultivate a particular doxastic state that is reassuring and confidence-building, he cannot respond directly in deliberation to considerations of that belief's usefulness. But as I shall argue, rationalization involves responding directly to considerations that thinkers know very well are irrelevant to the truth of the proposition under consideration.

In this paper I construct a model of rationalization that is designed to reconcile it with the exclusivity of doxastic deliberation. The model I propose involves the rationalizer engaging in a pretence (in the sense of "making as if"). I give reasons to think that only if we understand rationalizers as engaging in a form a performative pretense can we make sense of characteristic structural features of deliberative rationalization.

The argument proceeds as follows. First, I review Nishi Shah and J. David Velleman's influential account of "transparency". In section two, I present four ways misleading ways of understanding what the phenomenon of transparency amounts to, and I distill transparency into the core notion of "exclusivity". In section three, I present a paradigm of rationalization and analyze four key features. In section four, I present of model of rationalization as pretense that is consistent with a commitment to the exclusivity of doxastic deliberation. In the final section I describe how "out in the wild" the difference between rationalization and honest inquiry is often one of degree rather than of kind.

I. Transparency and the exclusivity of doxastic deliberation

Nishi Shah (2003) characterizes the special role that truth has in doxastic deliberation in terms of *transparency*, which he describes thus:

> [W]hen asking oneself whether to believe that p, [one] must ... immediately recognize that this question is settled by the answer to the question whether p is true[.] ... Within the perspective of first-personal doxastic deliberation, that is, deliberation about what to believe, one cannot separate the two questions. (447)

Shah re-iterates this view slightly differently in a later paper:

> To be clear, the feature that I call 'transparency' is this: the deliberative question whether to believe that *p* inevitably gives way to the factual question

> whether *p*, since the answer to the latter question will determine the answer to the former (Shah 2006: 481).

For Shah, transparency is an essential feature of "the phenomenology of deliberation." (2003, 462) Shah claims that from the perspective of the deliberating subject, the question whether to believe *p* is answered by, and only by, the answer to the question whether *p* is true. Truth excludes all other reasons for belief, and does so immediately. Moreover, a deliberating subject need not marshal any special effort to ignore non-truth directed considerations, because upon asking the question whether to believe that *p* her attention immediately focuses exclusively on the evidence. There is no "inferential step" from "Should I believe that *p*?" to "Is *p* true?" The question *whether to believe that p* "seems to collapse" into the question *whether p is true* (227).

Both Shah and his later co-author David Velleman (2005) acknowledge that belief is sometimes influenced by non-alethic factors such as wishful thinking (2005, 501). In fact, they take it as a virtue of their account that it "leaves room for the possibility that beliefs can be influenced by non-evidential considerations, because the view entails that one is forced to apply the standard of correctness *only in situations in which one exercises the concept of belief*." (501, my italics). The situations they have in mind are of course deliberative (ratiocinative) contexts.

Shah & Velleman contend that when one deliberates about whether to believe that *p*, this question "not only gives way to the question *whether p* but does so to the exclusion of any other, competing questions, such as whether *p* would be in one's interest." (501) The truth or falsity of *p* has "absolute priority" over all other considerations; it "crowds out" all competing non-epistemic questions. They content that any satisfactory account of belief must explain "the fact that truth occupies the sole focus of attention in doxastic deliberation." (2005, 500) Shah & Velleman insist that in order to explain transparency, "the degree of evidence-responsiveness required by the concept of belief would have to be such as to rule out other influences." (501)

It is important to note that "transparency" in Shah & Velleman's sense is a descriptive rather than a normative claim about doxastic deliberation. They do not maintain that in order to deliberation *well* or *rationally* or *effectively* one's deliberation must exhibit transparency (that would be to make the normative or prescriptive claim). Rather, they think that doxastic deliberation always in fact manifests transparency. This putative descriptive fact that "only truth regarding considerations move an agent in such deliberation" (468) is meant to be a deep, manifest, and inescapable feature of the phenomenology of deliberating about whether to believe that *p*.

Appealing to an inference to the best explanation, Shah & Velleman argue that transparency can be accounted for by positing that the *concept* of belief includes a standard of correctness.[6] When a subject deliberates about whether to believe that *p*, she exercises the concept of belief. And when she exercises the concept of belief, it is a "closed question" whether to believe that *p*, once it is ascertained that *p* is true. Shah & Velleman contend that the dispositions constitutive of possessing the concept of belief (and of seeking to answer a

[6] See Wedgwood 2002 for a separate defense of the conceptual claim that truth is belief's standard of correctness.

question framed with that concept) are responsible for the consciously felt authority of truth for belief in any deliberation that aims to settle belief.

It is not my aim in the paper to assess Shah & Velleman's normative account of the aim of belief. This task has been taken on competently and exhaustively by others and my quarry lies elsewhere. My aim is to refine the core notion of the phenomenon of transparency – which I refer to exclusivity – and to construct a model of rationalization that is consistent with it.

II. Precisifying Transparency: Exclusivity

In this section I identify four misleading ways of understanding transparency. This process of winnowing away inessential and misleading features will allow me to identify the core phenomenon, which, following Steglich-Petersen, I call "exclusivity".

First, the considerations adduced in deliberation need not *in fact* be truth-relevant. A person could mistakenly think that a consideration is truth relevant, and that consideration may still play a role in the deliberative formation of belief. For instance, in deliberating about whether in a vacuum a rock would fall noticeably faster than a feather, a person may think that the mass of the each of these objects is relevant to his deliberation and accordingly take this into consideration. Of course, he would be mistaken. But the mere fact that this consideration is not relevant does not preclude it from playing a role in the deliberative formation of belief. What is important for transparency is whether the thinker *regards* the consideration as relevant, not whether it is *in fact* relevant.

Second, we can and do explicitly consider truth-irrelevant considerations that we know to be truth-irrelevant in deliberation about what to believe. The transition between "Should I believe that *p*?" and "Is *p* true?" need not be frictionless or immediate. The phenomenology of doxastic deliberation does not bear out Shah & Velleman's contention that "truth occupies the sole focus of attention." (2005, 500) Nor does accord with Shah's later remark that "the question whether *p* is true *hegemonically* impose[s] itself on our doxastic deliberations." (2006, 488, italics in original) Indeed, the fact that one frames the deliberative question as "Should I *believe* that *p*?" rather than "Is *p* true?" is often an indicator that one has turned one's minds toward non-alethic considerations. Consider the person who thinks, "The balance of the evidence points to the treachery of my friend, but still, *should I believe this*?" Such a person might have in mind the implications for his relationship of the belief "my friend has betrayed me" or he might have in mind the implications for his own happiness. These are considerations that he is well aware are irrelevant to the truth of the proposition.[7]

Third, not only is it possible to consider non-truth-relevant considerations in deliberation about what to believe, but these sorts of considerations can and do influence the outcome of deliberation. An individual's view that in building his self-confidence for a job interview it would be useful to believe certain things about himself may well effect the body of evidence he examines as well as the evidential standards he applies. Researchers in psychology find

[7] I suspect that Shah and Velleman would deny that this person is engaged in doxastic deliberation when he turns his mind to questions such as these. This move requires that we have a conception of "pure" doxastic deliberation that is idealized from what we ordinarily understand as deliberation about what to believe "out in the wild". I see no problem with making this distinction in principle.

that when faced with an unpalatable proposition, people frame their inquiry with the question, "*Must* I believe this?" Contrariwise, when people are faced with a proposition that they wish for whatever reason to accept as true, they frame their inquiry with the question, "*Can* I believe this?" In the latter case, their permissive evidential standard is manifested in partial or truncated search for evidence (Dawson, Gilovich, & Regan, 2002; Ditto & Lopez, 1992; Ditto, Scepansky, Munro, Apanovich & Lockhart, 1998), consideration of a biased assemblage of evidence (Dunning, Meyerowitz, & Holzberg, 1989; Kunda, 1987; Lord, Ross, & Lepper,1979), and superficial processing of available information (Ditto et al., 1998). In the latter case, their stringent evidential standards are manifested by a relatively thorough search through all relevant information, maximizing the chances that any flaws or limitations of the data will be spotted (Dawson et al., 2002; Ditto et al., 1998; Ditto & Lopez, 1992).

Finally, we should not be tempted to think that the process of deliberation inoculates the deliberator from the influence of wishful thinking. In discussing the standard case of the self-deceived cuckolded husband, Shah contends that "if the husband turns his mind to the question whether to believe that his wife is faithful, then the concept of belief engages his thought, directing him to accept a proposition about his wife's fidelity only if he can discern its truth." (473) Contrariwise, Shah contends that if the husband "never bothers to ask himself this question […] then he may very well be induced by wishful thinking or other non-evidentially sensitive processes to be in a state of mind that, third-personally, we would judge to the belief that his wife is faithful." (473) But as we shall see there is no reason to think that wishful thinking is any less present in deliberative contexts than it is in non-deliberative contexts. To the contrary, there is much evidence from psychology that the process of deliberation provides occasion for particularly sophisticated and elaborate forms of wishful thinking. As I will argue in the next section, transparency and truth-regulation are entirely independent of each other. Transparency is a thesis about the *feeling* of ineluctable truth regulation, but it has no implications for *de facto* truth-regulation.

Having sought out and eliminated ways of understanding transparency that are not universal to the phenomenology of doxastic deliberation, we are now in a position to identify the core notion. The core notion is this: a thinker cannot in full consciousness decide whether to believe that *p* in a way that issues directly in forming a belief by adducing anything other than considerations that she regards as relevant to the truth of *p*. I will refer to this core notion as "exclusivity". Exclusivity is endorsed by nearly every philosophical camp, including pragmatists about reasons for belief. Richard Foley, who maintains that there are non-evidential reasons for belief, cites exclusivity to explain why we do not evaluate the rationality of our beliefs in terms of how well they promote our intellectual goals: "Offering you a million dollars to believe that the earth is flat may convince you that you have a good economic reason to believe the proposition, but in itself it won't be enough to persuade you that the earth is really flat." (1993, 16) For pragmatists like Foley, the psychological phenomenon of deliberative exclusivity has no normative implication despite its robustness, whereas for normativists like Shah & Velleman, the best explanation of this phenomenon is that the very concept of belief includes a standard of correctness. I do not try to adjudicate this disagreement here. It is the fact of exclusivity that is relevant to my investigation, not its explanation.

Even when narrowed and restricted in this way, deliberative exclusivity makes the phenomenon of rationalization look very puzzling. As I will argue, the rationalizer adduces

considerations that she *knows* to be irrelevant to truth of target proposition, she weighs these considerations quite explicitly in the guise of reasons, and she arrives at her conclusion directly by weighing these considerations. But before we can arrive at an understanding of how rationalization is possible in light of exclusivity, we need a clearer picture of what rationalization looks like. I turn to this in the next section.

III. A Picture of Rationalization

While contributors to the philosophical debate on the aim of belief make substantive claims about the phenomenology of doxastic deliberation, they rarely devote careful attention to its contours. In particular, in the discussion of transparency there is very little attention given to the distinctive manner in which non-evidential considerations find their way into otherwise ratiocinative belief-formation. I turn, then, to a literary example. A pitch-perfect rendering is to be found in the second chapter of Jane Austen's *Sense and Sensibility*. I analyze it in detail not just because it provides a brilliant illustration of the phenomenon in question, but also because it foregrounds the characteristic strategies of and constraints on rationalization, strategies and constraints that form the basis of my model.

Some quick background: When Mr Dashwood dies, his estate passes directly to his only son, John Dashwood. Mr Dashwood's second wife and their daughters, Elinor, Marianne, and Margaret, are left only a small income. On his deathbed, Dashwood extracts a promise from his son to use his inherited fortune to take care of his half-sisters. In what follows I reproduce in some detail a conversation between Mr. Dashwood (the son) and his wife Fanny, where they deliberate about exactly how much is owed to the half-sisters. Earlier, John Dashwood had decided that a lump sum of three-thousand pounds, as recommended by his father, would be the correct amount. He supposes that to set aside this amount of money for his half-sisters would be sufficient to secure their financial security and would discharge his obligation to his father . Dashwood's anticipation of his own substantial inheritance "warmed his heart, making him feel capable of generosity."[8] Over the course of a conversation with his wife, however, his magnanimous feelings give way to his wife's meanness. At the beginning of the conversation he needs to convinced by his wife; toward the end of the conversation he is a fully fledged co-rationalizer. Below I have excerpted and numbered the rationalizing considerations they adduce so that I can refer to them later.

Upon hearing of the three-thousand pounds, Austen's narrator describes Fanny Dashwood imploring her husband John:

(1) How could he answer it to himself to rob his child, his only child too, of so large a sum?

Fanny then casts doubt on the sisters' claim to the money, considering their pedigree:

(2) And what possible claim could the Mrs. Dashwoods, who were related to him only by half blood, which she considered no relationship at all, have on his generosity to so large an amount?

She also casts doubt on whether Dashwood the Elder was sound of mind when he extracted the promise form his son, John:

[8] The following quotations from *Sense and Sensibility* can be found on pages 5-10 of the 1833 Bentley edition freely available on Google Books.

(3) "He did not know what he was talking of, I dare say; ten to one but he was light headed at the time. Had he been in his right sense he could not have thought of such a thing as begging you to give away half your fortune from your own child."

John Dashwood begins to allow himself to carried along:

(4) "He did not stipulate any particular sum, my dear Fanny, he only requested me in general terms, to assist them, and make their situation more comfortable than it was in his power to do."

Intuiting his wife's unwholesome purposes and hesitating , John Dashwood insists that he will not break his promise to his father: "The promise, therefore, was given, and must be performed."

But Fanny Dashwood sees, in the vagueness of the content of the promise, an opening:

(5) "Well, then, let something be done for them; but that something need not be three thousand pounds. Consider […] that when the money is once parted with, it can never return. Your sisters will marry and it will gone forever. If, indeed, it could ever be restored to our poor little boy –"

John Dashwood accepts the invitation to fix upon his son's welfare:

(7) "Why to be sure […] that would make a great difference. The time may come when Harry might regret that so large a sum was parted with. If he should have a numerous family, for instance, it would be a very convenient addition."

Responding to his wife's entreaties, John reduces the sisters' inheritance by half:

(8) Perhaps, then, it would be better for all parties if the sum were diminished by one half. – Five hundred pounds would be a prodigious increase to their fortunes.

Remarking that the sisters could hardly expect more, Fanny Dashwood retorts:

(9) "There is no knowing what they may expect […] but we are not to think of their expectations: the question is, what you can afford to do."

John Dashwood recalls the fact that the young women are also heirs of their mother:

(10) "Certainly; and I think I may afford to give them five hundred pounds apiece. As it is, without any addition of mine, they will each have about three thousand pounds on their mother's death- a very comfortable fortune for any young woman."

John Dashwood has the idea of giving the money to their mother instead, not all at once but as a modest annuity:

(12) "I do not know whether, upon the whole, it would not be more advisable to do something for their mother while she lives, rather than for them- something of the annuity kind I mean. My sisters would feel the good effects of it as well as herself. A hundred a year would make them all perfectly comfortable."

Fanny Dashwood hesitates to give her assent to the plan:

(13) "But, then, if Mrs. Dashwood should live fifteen years, we shall be completely taken in."

When John Dashwood points out how unlikely the sisters' mother will live that long, Fanny Dashwood brings up the fact that her mother's servants, who were given an annuity, lived a very long time:

(14) "but if you observe, people always live forever when there is an annuity to be paid them."

John Dashwood is only too happy to be persuaded:

(15) It is certainly an unpleasant thing […] to have those kind of yearly drains on one's income. One's fortune, as your mother justly says, is not one's own. To be tied down to the regular payment of such a sum, on every rent-day, is by no means desirable: it takes away one's independence."

Fanny Dashwood points out another disadvantage of the annuity is that it will occasion no sense of gratitude on the part of the sisters:

(15) [A]nd, after all, you have no thanks for it. They think themselves secure; you do no more than what is expected, and it raises no gratitude at all.

She also emphasizes that their own financial security is never certain:

(16) "It may be very inconvenient some years to spare a hundred, or even fifty pounds from our own expenses."

By now an equal participant in their joint rationalization, John Dashwood decides that the sisters should get only "fifty pounds now and then" to ensure that they do no fritter away their money:

(17) I believe you are right, my love; it will be better that there should by no annuity in the case: whatever I may give them occasionally will be of far greater assistance than a yearly allowance, because they would only enlarge their style of living if they felt sure of a larger income, and would not be sixpence the richer for it at the end of the year. It will certainly be much the best way. A present of fifty pounds, now and then, will prevent their ever being distressed for money, and will, I think, be amply discharging my promise to my father."

John Dashwood emphasizes how few the sisters' expenses may be:

(18) "They will live so cheap! Their house-keeping will be nothing at all. They will have no carriage, no horses, and hardly any servants; they will keep no company, and can have no expenses of any kind! Only conceive how comfortable they will be! Five hundred a year! I am sure I cannot imagine how they will spend half of it; and as to your giving them more, it is quite absurd to think of it."

As if to disperse any lingering doubts, Fanny Dashwood declares :

(18) " And I must say this, that you owe no particular gratitude to him, nor attention to his wishes; for we very well know that if he could, he would have left almost everything in the world to them."

The narrator concludes:

(19) This argument was irresistible. It gave to his intentions whatever of decision was wanting before; and he finally resolved, that it would be absolutely unnecessary, if not highly indecorous, to do more for the widow and children of his father than such kind of neighborly acts as his own wife pointed out.

There are number of things that can be learned from Austen's rendering of this couple's undeniably impressive rationalization. The manner in which the Dashwoods arrive at their conclusion is rather subtle and merits close attention. To get a clearer focus on what they are doing, it will be useful to contrast this case of deliberative rationalization with four other related, but distinct, activities.

It's important to note that neither John Dashwood nor Fanny Dashwood explicitly represents to the other or to themselves that the goal of their deliberation is to reach the conclusion that what is owed to the sisters is far less than the three-thousand pounds initially supposed. The strategy of the Dashwoods involves a kind of misdirection that is absent in deliberation of the person who knowingly and in full consciousness aims to cultivate a state of mind that he hopes will conduce to belief in a particular proposition. Consider, for example. the contrast case of a composer suffering from crippling self-doubt who on the advice of his therapist embarks on a project of self-affirmation, deliberately and strategically directing his attention at regular intervals to hallmark past accomplishments each time his belief in his own capacity to compose is clouded with uncertainty. For this man, the content of the target belief ("I am a competent composer") is presented clearly and explicitly, as are the considerations in support of the target belief ("I composed an excellent fugue last year. I composed a superb sonata last month…")

The rationalizing of Austen's characters is also distinct from the strategy of the person who tries to bring himself to believe in God by applying the logic of Pascal's Wager. In trying to bet his way to theistic credence, this person explicitly presents himself with a non-epistemic reason to believe, and he presents that reason *as* non-epistemic. He tries to combine the aim – believe *p* only if *p* is true - with a pragmatic aim– believe *p* because it makes salvation more likely and damnation less likely.[9] Pascal himself saw the difficulty with this approach, which is why he also counseled surrounding oneself with faithful people, going to mass, and taking holy water.[10]

While the Dashwoods do not try to weigh evidential reasons against pragmatic ones, neither do they deploy the kind of indirect tactics that Pascal suggests. It is not their strategy to cultivate behavioral habits or immerse themselves in a social environment conducive to the formation of a particular belief. Granted, they may well avoid the sorts of people they suspect would prick their consciences or debunk their self-justifications. But these later evasive maneuvers are not essential to the manner in which they reach their conclusion.

Finally, what happens in Austen's dialogue is also distinct from what happens in the shopworn philosophical example of the self-deceived cuckold who comes to believe in his wife's fidelity because the thought of her infidelity is so abhorrent to him. While such a case does involve an unarticulated refusal to face up to readily available evidence, it does not involve (at least as it is usually described) the careful and protracted weighing of considerations. We picture the man's desire that he not be a cuckold as a irrational "influence" on his beliefs, a blind force that gums up the works of rationality in the way that

[9] David Owens (2003) argues that it is impossible to weigh the "aim of belief" – believe p only if p is true – with other aims. This leads him to think that there is no aim of belief at all (since aims are the sort of things that we can weigh).

[10] Pensée #233. Gérard Ferreyrolles, ed. Paris: Librairie Générale Française, 2000.

a strong magnet impedes the proper functioning of a computer. The man's desire does not present itself in the guise of reasons nor does he work with the evidence in the manner of a deliberating agent.

It is noteworthy that Mr. and Mrs Dashwood *do* work with the evidence. The considerations that they adduce are not the product of idle fantasy or pure whimsy. On the contrary, on the route to their deliberative conclusion the Dashwoods take into consideration a number of verifiable facts and plausible conjectures. Take consideration #1: passing on more money to the sisters might have the effect of making their son less well off. This is reasonable. Fanny's insistence that the Dashwood sisters are related only by half blood (consideration #2) is also firmly rooted in reality. And it is also the case the Dashwood the Elder did not stipulate a specific sum for the half-sisters (consideration #4).

Not only do many of the considerations adduced the Dashwoods appear to be constrained by the truth, these considerations play an essential role in the deliberative formation of their conclusion. They do not arrive at their conclusion "arbitrarily" or "at will". Indeed, the Dashwoods have in common with honest inquirers that there is never any guarantee that they will reach a conclusion that is amenable or desirable. Honest inquirers risk arriving at conclusions that are unwelcome because such conclusions may not be best supported by the evidence. The Dashwoods risk arriving at conclusions that are unwelcome because it takes great care and ingenuity to construct a rationalization that issues in the desired conclusion and that at the same time resists easy and obvious debunking.

The Dashwoods can and do acknowledge the considerations that lead them to their conclusion: they articulate these considerations explicitly. In a key footnote, Shah & Velleman (2005) address this worry: "Our claim here is not that deliberation about what to believe cannot be influenced by non-evidential considerations; it is that such deliberation cannot explicitly treat such considerations as relevant to the question what to believe. Any influence that such considerations exert must be unacknowledged." (footnote 16) In a later paper, Shah (2006) later reiterates and elaborates this idea:

> This does not mean that deliberation about what to believe cannot be influenced by non-evidential considerations; it just means that such deliberation cannot explicitly treat such considerations as relevant to the question of what to believe. Any influence that such considerations exert must be unacknowledged. But this is just as it should be. Transparency is a conscious phenomenon: we cannot consciously acknowledge considerations which are irrelevant to the truth of *p* as determining whether to believe that *p*. However, as we know every well, what cannot be consciously acknowledged often has a powerful influence none the less. (489)

Attending to the phenomenology of rationalization forces us to revisit Shah & Velleman's contention that the influence that non-epistemic considerations exert must be "unacknowledged". As our example shows, the rationalizers need not avoid acknowledging what their reasons are: their conclusion survives this acknowledgement. What they cannot acknowledge is the fact that the considerations they adduce are non-conclusive or even irrelevant to establishing the truth of their conclusion. For the rationalization to remain intact, rationalizers must forswear conscious acknowledgement of certain *features* of the deliberative considerations, not the considerations *themselves*. This requirement is distinctive

of first-personal deliberation. Consider the contrasting mental attitudes of Austen the author, who in composing the dialogue may present to herself quite clearly the flimsy and self-serving quality of the reasoning of her characters, and that of the Dashwoods, who cannot attend to these qualities if they are to arrive at their conclusion.

But given the assumption that Dashwoods are no dolts (especially not Fanny Dashwood!), it is easily within their ken to grasp the flimsiness of their reasoning. Indeed, the moral intuition that they are blameworthy for such self-serving reasoning (Austen clearly means for us to feel this way) is incompatible with thinking that they are excused by naïveté or feeble-mindedness. If we think the Dashwoods are simply *deluded*, the moral opprobrium with which we regard them would not make sense. And although Fanny may be the instigator, John Dashwood is not to be exonerated since their joint rationalization would fail in the absence of mutual cooperation. Here we find a disanalogy between joint rationalization and joint honest inquiry: the former requires that each individual's ulterior motivations are aligned with those of the others, whereas the latter requires a mutual commitment to follow where the evidence leads.

Rationalizers employ a distinctive set of strategies to reach their desired conclusions. For instance, they may adduce pseudo-reasons, considerations that have only the appearance of relevance to the deliberative question. Consideration #15 and #18 are of this type. It is of not relevant to the question of what the sisters are due whether they can live cheaply or whether they will feel gratitude toward their "benefactor". Also consideration #16 – just because Mr. Dashwood owes no debt of gratitude to Dashwood the Elder does not imply that he is released from his promise. Rationalizers may also adduce weak reasons, considerations that are relevant to the question at hand, but that are given undue weight or taken as conclusive reasons. Consideration #1 and #3 are of this type. Of course the Dashwoods must consider the wellbeing of their son, but giving the half-sisters a decent inheritance is unlikely to imperil him. Finally, rationalizers will often support their conclusions with empirical claims that are difficult to verify or to falsify. See here the conjecture that people who receive an annuity live longer (consideration #14), or the conjecture that in the future it may be painful for their family to spare even fifty pounds (consideration #16).

Although this taxonomy is probably not exhaustive, I think that it is illustrative. What all of these strategies have in common is that they inculcate the *appearance* of sound reasoning while still affording crucial flexibility in the outcome of deliberation. Notice that the transparency phenomenon (the phenomenon whereby the deliberative question gives way to the factual question) is just as robust in rationalization as it is in honest doxastic deliberation. In both cases, the discursive move *p is true* ultimately settles the question of *whether to believe that p*. Reaching the conclusion *p is true* is just as decisive for the rationalizer as it is for the honest inquirer. In the pragmatics of deliberation, honest or not, this kind of discursive move serves as a "deliberation-stopper". Shah & Velleman characterize this transition as immediate, effortless, and ineluctable. In Section 2 I gave reasons to be skeptical of the universality of the smoothness and immediacy of the transition from the alethic question to the doxastic question. Nonetheless, to the extent that this transition ultimately is made in the deliberative formation of belief, it is no less the case in episodes of rationalization. Were the Dashwoods to agree that, "It is true that we owe the sisters nothing more than kind and neighborly acts" the question of whether to believe this proposition would be settled.

Accordingly, the relative strength or weakness of truth-regulation in the deliberative formation of belief and the feature of transparency are entirely independent of each other.

III. A pretense model of rationalization

A satisfactory model of rationalization will reconcile these three features:

(1) *Deliberative exclusivity:* A thinker cannot in full consciousness decide whether to believe that *p* in a way that issues directly in forming a belief by adducing anything other than considerations that she regards as relevant to the truth of *p*.
(2) *Non-naiveté:* Rationalizers know that the considerations they adduce are not sufficient to establish the conclusions they reach.
(3) *Deliberative weighing:* The considerations adduced in the process of rationalization *do* play an essential role in the deliberative formation of the conclusion. The influence of the rationalizing consideration on belief goes via the content of the consideration, and the consideration is weighed explicitly.

In order to reconcile the fact that rationalizers *know* that the considerations they adduce do not establish the conclusions they reach with the fact that they cannot so *regard* them, I will argue that we must understand what rationalizers do as engaging in performative pretence. While it is not possible for an individual to believe that *p* when she knows that *p* is false, there is no difficulty with pretending that *p* (in the sense of making as if *p*) when she knows that *p* is false.

In advancing a pretense account of rationalization I must address the understanding of pretense as off-line processing that is altogether segregated from belief and devoid of motivational force. Nichols and Stich (2000), for example, explicitly set out to provide an model of pretense that explains how it is that "the events that occurred in the context of pretense have only quite limited effect on the post-pretense cognitive state of pretender." (120) While this "quarantining" of pretense contents from belief contents is typical, it is not universal. As Gendler (2003) points out, in some contexts quarantining gives ways to its opposite – "contagion" – whereby the pretended contents come to be believed, or treated as if they are believed, merely because they are pretended. In cases of affective transmission mere contemplation of a content that is emotionally charged causes the thinker to behave and feel in such a way that is consistent with belief in that content (131). This explains why if Mr. Dashwood is later accused of reneging on his promise he may respond with genuinely felt outrage and indignation.

In order to understand rationalization as performative pretense, we must also address the understanding of pretense whereby the pretender processes belief-eligible content in same way that she processes belief. For example, the Nichols and Stitch model of pretense sets out to explain how it is that "inference mechanisms treat pretense representations in roughly the same way that the mechanisms treat real beliefs." (125) Although Nichols and Stitch are right that such "mirroring" is typical, Gendler points out that pretense episodes may also manifest "disparity", the tendency whereby pretense content differs from non-defective belief content in that what is pretended may be incomplete (some features may remain

permanently unspecified and unspecifiable) as well as incoherent (some features may be logically and conceptually incompatible) (137).

In the light of this more nuanced understanding of the relationship between belief and pretense, it is useful to re-examine Shah's discussion of the self-deceived cuckold. Shah maintains that, "if he views the attitude as a belief, then he will take evidence to be solely relevant, but if he thinks of the attitude as something he is assuming for the sake of a *pretense*, for example, then he won't take evidence to bear at all on whether to maintain or abandon the attitude." (2003, 468) The Dashwood rationalization subverts this kind of dichotomous analysis. While the Dashwoods clearly do not proceed in the same way they would if they took the evidence to be solely relevant to their deliberative question (witness their low evidential standards and irrelevant considerations), they are in important ways constrained by the evidence (witness their eschewal of outright contradiction and manifest falsity). What best explains the complex structure of their rationalization? I will argue that the best explanation is that they are engaged in a kind of pretense whose content includes their being guided by the aim of inquiry, which requires that the conclusion they reach is correct only if there is sufficient reason to believe it. Accordingly, rationalizers *make as if* the considerations they adduce are sufficient to establish the conclusions they reach, even though they know that this is not the case. Their episode of pretense exhibits mirroring in that it reproduces the rhetoric and characteristic discursive moves of honest inquiry. At the same time it also manifests disparity in its tolerance for (suitably disguised) incoherent content.

As with the case of explicit fictions, being "realistic" helps with the suspension of disbelief. The story told cannot be wildly implausible, far-fetched, or manifestly self-contradictory. If it is, the rationalization will be unstable and vulnerable to easy debunking. Just as reader of a novel may "pop out" of story whose plot is obviously incoherent, so also a thinker will not be moved by a rationalization that lacks the basic characteristic discursive moves of honest inquiry. The ability to rationalize is an achievement requiring a good deal of cognitive sophistication and individual thinkers will vary in their possession of the skill required.[11] The journeyman rationalizer will concoct rationalizations that are vulnerable to easy debunking by his own conscience or by the criticism of others. The maestro rationalizer will concoct far ranging rationalizations that ultimately efface any awareness of their falsity.[12]

Here one might object by asking whether it is plausible to suppose rationalizers are engaging in pretense. The experience of rationalizing certainly *feels* rather different from acting in a play or participating a child's game of make-believe. Surely a rationalizer is not consciously thinking of the attitude he is adopting as something other than belief? But just as a method actor who is preparing for a role can pretend that certain things are true of his life without consciously attending to the fact that this is what he is up to, so too the rationalizer may pretend that certain considerations provide conclusive reasons for belief without occurent conscious awareness of the fact that he is pretending. The difference between the rationalizer and the method actor is that the method actor has at an earlier time consciously

[11] Ziva Kunda (1990) provides an excellent survey of classic psychological work on the mechanisms subserving motivated reasoning (strategies for accessing, constructing, and evaluating beliefs) that suggest that subjects exhibiting motivated reasoning are constrained by their variable ability to produce seemingly reasonable justifications for their conclusions.

[12] John LeCarré's novel *Little Drummer Girl*, the story of an English actress turned double-agent, is a masterful illustration of this phenomenon.

framed his activity as one of imaginative and performative pretense, whereas the rationalizer has not.

Often it is only in abeyance that a individual becomes fully aware that he has been adopting the projective attitude of pretence rather than the receptive attitude of belief. Consider, for example, the undergraduate student who is exploring new fields and trying to discover where his passion resides. He sits in a freshman seminar on "Althusser and Death" with rapt attention (although he understands little of what is said), emulating the stylish sophistication and attractive mysteriousness of his classmates. It is only in later years when he gets a taste of what it really feels like to be passionately immersed in a subject he loves that he realizes that in fact he was pretending to be interested in Althusser all this time. If he were able to understand himself better as a freshman, he may have realized this more quickly. (It should be evident that a case like this is importantly different from a case of a cynical student who consciously decides to adopt the tactic of pretending to be interested in Althusser in order to, say, impress his instructor). A rationalizing deliberator does not "think of the attitude as something he is assuming for the sake of a *pretense*", at least not at the time of deliberation. Rather, he "slips into pretending" without attending to the fact that he is doing this.

It can be particularly hard for an individual to keep track of belief and pretense in contexts where it is unclear to her whether others are pretending, as in case of "open lies". Consider the case of Meredith, who after three and a half months of pregnancy has a "baby bump" that is beginning to show through her loose fitting clothes. She has not yet announced to her friends and colleagues that she is pregnant, and she is unsure of whether they are pretending not to know, or whether they are in fact ignorant. After a while, she herself loses track of whether *she* is pretending they do not know, or whether she believes this. When she finally makes the announcement, their expressions of astonishment seem somewhat forced, and Meredith finds herself unsurprised at their affectation. It is only at this point that she realizes that she was only pretending (rather than believing) to be unaware that they were pretending not to know. (It should be evident that a case like this is importantly different from a case where Meredith assigns a low degree of credence to her belief that her friends and neighbors are pretending in a way that is altogether self-aware).[13]

Readers might not be convinced that one can pretend without knowing that one pretends from these admittedly merely suggestive examples whose interpretation may be contested. One might object that it does violence to the concept of pretense to admit cases where one is not occurently aware that one is pretending, and where one has not earlier framed one's activity as an episode of pretense (as one has in the case of method acting). On this matter I am content to cede ground. It is not essential to my account that "pretending" in

[13] Paul Griffiths and Andrea Scarantino (2009) have undertaken some very suggestive empirical work on sulking that seems to indicate that sulking behaviors are modulated by strategic aims of which the subject is largely unaware at the time of their expression. Although the sulking behavior is often in part performative pretense, subjects do not seem to be occurently aware of this fact.
J. David Velleman (2000) analyzes an entire category of "arational" actions as instances of pretending where you not presently think of yourself as pretending. Examples include talking to yourself where you are imagining yourself in conversation with someone else, saying things that you wish you had said or could say (264). He maintains that such actions can only be explained in terms of "wishes" and "imaginings" (rather than beliefs and desires). But in paradigm cases it is only abeyance that we become aware that we were pretending in this way.

rationalization is cut from the same cloth as the pretending that we do in imaginary games or dramatic acting. It is the features of the imagination - mirroring and disparity, quarantining and contagion - that are essential for my model of rationalization. If readers prefer to distinguish pretence from "pretense-like rationalization" (schmetense, if you like) I have no objection.

4. Rationalization and Honest Inquiry

A common feature of everyday doxastic deliberation is that practical and epistemic considerations are often woven very fine. Consider a married couple's deliberation about whose elderly parents are more feeble, where the veiled subtext is a negotiation about whose parents they ought to visit next. Each side advances considerations that emphasize the frailty of his or her parents (and therefore the necessity of a visit): cataracts, a gouty foot, an irregular heartbeat. In the end, there may be a consideration adduced by one side that is so weighty and so manifest that it definitively settles the question (say, the urgent need for bypass surgery). If this is the case, the rationalization of the other side will be quickly short-circuited. But this does not imply that the considerations on the other side were totally "made up" or "arbitrary". The illnesses need not have been invented.

Recall Gendler's way of distinguishing belief from pretense: "If I bear an attitude of belief toward P, I should be willing to submit my evidence for P to rational scrutiny, and I should be committed to abandoning my belief if I acquire grounds for thinking it false. […] If I bear an attitude of pretense towards P, I am not committed to submitting my evidence for P (should I have any) to rational scrutiny, nor am I committed to abandoning my pretense if I have or acquire grounds for thinking it false." (237-8) Attending closely to the phenomenology of doxastic deliberation shows that this dichotomy is not as clear cut as it may seem. The "commitment" in question is not "off" or "on": it may vary by degree, and it may be stronger or weaker with respect to different deliberative considerations. In contexts of rationalization, unlike standard cases of self-aware imaginative pretense, the subject often submits his evidence to a moderate degree of rational scrutiny (although, of course, not scrutiny that is so intense that the rationalization will crumble). Conversely, should the rationalizer become vividly aware of grounds for thinking that his conclusion is false, he will be forced to abandon it.

Deliberative episodes may vary in their relative proximity to the pole of "shameless rationalization" and that of "unadulterated inquiry". A subject may exercise a good deal of care and scrutiny in his investigation, but still steer clear of certain narrowly defined hypotheses that might lead to conclusions that are "unthinkable". Consider the case of a detective who doggedly and exhaustively follows all his leads except for those that may bring him to believe in guilt of his son. At the other end of the spectrum, a subject may require only the flimsiest cover to adopt what is evident to everyone else as a very implausible posture. (A student in my ethics class once told me that she thought vegetarianism was immoral because if humans ate all the plants nonhuman animals would starve to death.)

Rationalization is marked by an inherent tension between arriving at the desired conclusion (e.g. - I am honest, I am rational, I am good) and getting there with a plausible story. The negotiations of these aims is an at time protracted balancing act that can be performed with

varying degrees of proficiency. It is never guaranteed at the outset that rationalization will realize its aims. While rationalizers are partially constrained by their evidence base, they do their utmost to relax the restriction. Sometimes the goal is defined positively (e.g. – arriving the conclusion that one is faithful to a promise). Sometimes the goal is defined negatively (e.g. – avoiding the conclusion that one's grandson is callous). In paradigmatic cases, rationalization is something that cognizers *do* rather than something that befalls them: it is not a mere "influence" on their belief-forming processes. In contexts where facing up to the truth is morally weighted or practically important, the rationalizing deliberator is guilty of a kind of culpable negligence. His failing is that he does not attend to the fact that he pretends. The severity with which we view this recklessness will depend on topic and the purpose of the rationalization. Buttressing one's own self-confidence is one thing, cheating someone of an inheritance is another. But the contours of the moral evaluation of rationalization is a topic for another investigation. This paper lays the foundation for that investigation by giving an account of just what it is that rationalizers do.

Works Cited

Boghossian, Paul A. 2003. "The normativity of content" *Philosophical Issues* 13: 31-45..

Dawson, Gilovitch, and Regan. 2002. "Motivated Reasoning and Performance on the Wason Selection Task" *Personality and Social Psychology Bulletin* 28:10, 1379-1387.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, 63, 568-584.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovich, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, 75, 53-69.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. Journal of Personality and Social Psychology, 57, 1082-1090.

Elga, Adam (2005). "On overrating oneself. . . and knowing it." Philosophical Studies 123: 115–124.

Festinger, L. 1957. *A theory of cognitive dissonance.* Stanford: Stanford University Press

Foley, R. 1993. *Working without a Net* Oxford University Press.
Frankfurt, Harry. "On Bullshit." *Raritan Quarterly Review* 6, no. 2 (Fall 1986)

Gendler, Tamar. 2003. "On the Relation between Pretense and Belief." *Imagination, Philosophy and the Arts.* (Domenic Mclver Lopes & Matthew Kieran, eds.) NY: Routledge, pp. 125–141.

Gendler, Tamar. 2008. "Self–Deception as Pretense," *Philosophical Perspectives: Philosophy of Mind*

Glüer, K. and Å. Wikforss. 2009. "The Truth Norm and Guidance: a Reply to Steglich-Petersen" Mind 119: 757-761

Kunda. Z. 1990 "The Case for Motivated Reasoning" *Psychological Bulletin Vol. 108, No. 3, 480-498.*

Greenwald, A.G. & Ronis, D.L. 1980. Twenty years of cognitive dissonance: Case study of the evolution of a theory. *Psychological Review*, 85, 53-57.

Griffiths, Paul, and Scarantino, Andrea. 2009. "Emotions in the Wild: The Situated Perspective on Emotion," in *The Cambridge Handbook of Situated Emotion*, edited by Philip Robbins and Murat Aydede, 437-453. New York: Cambridge University Press.

Halzlett, A. 2013. *A Luxury of the Understanding: On the Value of True Belief* Oxford.

Kunda, Z. (1987). "Motivated inference: Self-serving generation and evaluation of causal theories." *Journal of Personality and Social Psychology*, 53, 636-647.

Kunda, Ziva. 1990. "The Case for Motivated Irrationality." *Psychological Bulletin* Vol. 108, No. 3, 480-498.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." *Journal of Personality and Social Psychology*, 37, 2098-2109.

Nichols and Stitch. 2000. "A Cognitive Theory of Pretense" *Cognition* 74: 115-147.

Pascal, B. *Les Pensées* Gérard Ferreyrolles, ed. Paris: Librairie Générale Française, 2000

Owens, D. "Does Belief Have an Aim?" *Philosophical Studies*. Vol 15: 3, 283-305

Shah, N. 2003. "How Truth Governs Belief" *The Philosophical Review* vol 112:4 447-482.

Shah, J. and J. D. Velleman. "Doxastic Deliberation" *The Philosophical Review* vol 114, No 4 497-534.

Shah, N. 2006. "A New Argument for Evidentialism." *The Philosophical Quarterly* vol. 56, No. 225.

Steglich-Petersen, A. 2006 "No norm needed: on the aim of belief" *The Philosophical Quarterly*, 56: 499-516.

Steglich-Petersen, A. 2009. Weighing the aim of belief. *Philosophical Studies*, Vol. 145, p. 395-405.

Wedgwood, R. 2013.. Doxastic Correctness. *Aristotelian Society Supplementary Volume* 87 (1):217-234.