

# Reflections on the Ideology of Reasons:<sup>1</sup>

John Hawthorne and Ofra Magidor

## §1. Setup

In this paper we offer a series reflections on the rather complex ideology of reasons. We will start in this section by spelling out some background assumptions we make. As we shall see, among the normative reasons for an agent  $X$  to  $\phi$ , it is common to distinguish between those reasons that the agent possesses and those which she does not. In §2, we argue that possession requires knowledge (if  $p$  is a reason for  $X$  to  $\phi$  then  $X$  possesses  $p$  as a reason only if  $X$  knows that  $p$ ). In §3, we argue, first, that the normative reason construction is factive (if the proposition that  $p$  is a reason for  $X$  to  $\phi$ , then it is true that  $p$ ), and second that possession ascriptions can be factored into a normative reason construction and a possession claim (the fact that  $p$  is a reason which  $X$  possesses to  $\phi$  iff the fact that  $p$  is a reason for  $X$  to  $\phi$  and  $X$  possesses this fact as a reason).

One important theme that runs through both sections is the following: there is a typical range of cases where, since an agent does not know a pertinent worldly fact that might otherwise serve as a motivating reason, one might be tempted to fall-back on describing the agent's reasons using a psychological ascription (e.g., in the case where an agent is hallucinating a tiger and runs, we might revert to 'His reason for running was that he thought there was a tiger in the room'). We maintain that in many such cases the psychological fact cited is not after all a motivating reason (though the ascription sentence might still have a true reading and the psychological fact might still count as a possessed normative reason). Indeed, we argue that in cases of this sort the agent often acts for *no* motivating reason.

In §4, we turn to compare two prominent views concerning the nature of normative reasons: Kearns and Star's view of reasons as evidence that one ought to  $\phi$ , and John Broome's view of reasons as explanations for why one ought to  $\phi$ . While both views have significant merit, we argue that they also face some non-trivial challenges, and discuss a range of considerations that can help to adjudicate between these two conceptions of reasons.

Let us begin, however, by noting some things we shall be taking for granted in the discussion that follows. First, we take it that according to the primary use of 'reason', reasons are propositions. This is not to deny that there are uses of 'reason' which cite things other than propositions as reasons (for example: 'That idiot was the reason for my leaving the party' or 'The stock crash was the reason for his getting depressed'). But it is overwhelmingly plausible that such constructions are derivative (from

---

<sup>1</sup> Thanks to audiences in LSE, Oxford, and in the central APA (especially so to Mark Schroeder) for helpful discussion. Thanks also to Michael Brady, John Broome, Christina Dietz, Julien Dutant, David Enoch, Maria Lasonen-Aarnio, Daniel Star, and Tim Williamson for their comments and suggestions.

truths like ‘The fact that that idiot was there was the reason for my leaving the party’ and ‘That the stock crash occurred was the reason for his getting depressed’).<sup>2</sup>

Second, we assume that the normative construction ‘Reason to phi’ is context sensitive across various parameters. (Note that it is unsurprising that this infinitival construction is normative: infinitivals are often normative. Consider, for example, ‘The thing to do is phi’.) Such constructions can encode either an ‘ought’ or a ‘may’ modality. And just as ‘ought’/‘may’ can take on a variety of meanings, so can the associated infinitival ‘reason to phi’ construction.<sup>3</sup> There are a variety of candidate meanings for ‘ought’ and ‘may’, some more subjective than others. We will work for the most part with the (radically oversimplified) picture according to which there are two ‘oughts’: a subjective and an objective one. We hope the oversimplification won’t matter much for what we have to say. Roughly our envisaged objective ‘ought’ ranks actions according to the best outcome, while the subjective ‘ought’ ranks according to the best expected outcome by the lights of the agent’s evidence.<sup>4</sup> (We think of evidence as what the subject knows, though much of what we say could be adapted to other frameworks for thinking about evidence.) This basic structure covers both reasons to act and reasons to believe.<sup>5</sup>

Third we assume the following taxonomy for classifying the various uses of the ideology of reasons. (While there is nothing like a universally accepted taxonomy in this area, what follows is not particularly idiosyncratic.) Quite apart from (i) normative reasons just discussed (i.e. reasons there are for an agent to act, believe, or feel a certain way) there are also (ii) explanatory reasons – reasons why the agent acted, believed or felt a certain way, and (iii) motivating reasons, the reasons for which the agent acted on a particular occasion. (Motivating reasons are called ‘personal reasons’ by Grice because we standardly describe them using possessive constructions of the form ‘X’s reason(s) for phi-ing was (were) that...’). Within the category of normative reasons, there is a special subcategory of possessed normative reasons – reasons to act that the agent possesses. The contrast between possessed and unpossessed reasons we have in mind is fairly intuitive. When a glass contains poison but an agent is unaware of this, there is a reason for the agent to avoid drinking from the glass, but that reason for avoidance is something that the agent is not in a position to use as a consideration when acting.

---

<sup>2</sup> We shall not undertake to argue here against the main rival, namely that states and events are reasons (in the primary sense). We also note in passing one interesting construction that we shall not focus on, namely, ‘Her/his reason for v-ing was to phi’. It is plausible that this is a propositional use - perhaps it is elliptical for ‘His/her reason for x-ing was that his/her goal was to phi’.

<sup>3</sup> One natural way of thinking about this is Kratzer’s (see e.g. Kratzer (1981)), where the various uses of ‘ought’ have a common skeletal logical structure but are distinguished by being associated with different ways of ranking states of affairs (she calls these modes of ranking ‘ordering sources’) and/or differing domains of states of affairs (‘modal bases’).

<sup>4</sup> Obviously there are also sorts of further distinctions (e.g. what is best relative to the agent’s goals or instead relative to some ethical or legal way of ranking states of affairs).

<sup>5</sup> If one wants to focus on a particular subset of reasons to believe – *epistemic* reasons to believe, one can do so by appealing to an epistemic kind of ordering-source for resolving the context sensitivity of ‘X ought to believe p’. (For example, we can rank beliefs as better or worse according to their instrumental goodness vis-à-vis the agent’s happiness, or instead as closeness to some epistemic ideal.)

The literature sometimes seems to suggest that we can distinguish whether possessed normative reasons are in play using superficial linguistic tests. We should be cautious here. First, possessive constructions are notoriously context sensitive. We think that they don't always indicate possessed normative reasons in the intended sense of 'possessed'. For example, there is a natural use of the possessive which allows us to say 'The agent has a reason to leave the building' when there is a bomb in the building, but the agent is completely unaware of this fact. But this use of 'has a reason' doesn't mark the kind of possession that theorists in the area are typically interested in.<sup>6</sup> With Mark Schroeder, we agree that a good indicator of whether a normative reason is possessed (in the relevant sense) is whether that reason is available as a motivating reason. Thus, even when our central concern in this paper is with possessed and unpossessed normative reasons we will often consider whether a certain proposition can function as a motivating reason, since this is a good heuristic for testing the kind of possession of interest.<sup>7</sup>

Finally, we should note that as we are thinking about it, one needs to keep apart several distinctions that are often conflated in the literature. First, a particular normative reason might be possessed or unpossessed whether or not it is reported as being possessed. Thus, suppose for example that the following report is true: 'There is a reason for X to flee the building, namely that the building is burning'. The report in itself does not commit us to the claim that the reason is possessed (X might be completely unaware that the building is burning), but neither does it rule out the possibility that this is a possessed reason (the report can be true even if X is fully aware that the building is burning). Second, the question whether the normative force of the infinitival construction is objective or subjective is orthogonal to the question of whether the reason in question is reported as being possessed. Thus, for example 'There is a reason for X to phi, namely that p', which does not report possession can be read as utilising either a subjective or objective normative ordering source (and the same is true of ascriptions which do report possession).

## **§2. Possession requires knowledge**

### **§2.1 The initial case in favour of possession requires knowledge**

What does it take for an agent to possess a reason? We claim that if the proposition that p is a reason for S to phi, then S possesses it as a reason to phi only if S knows that p.<sup>8</sup>

Given that in order for the proposition that p to be a reason it must be true that p (see §3 below), it is clear that there will be no cases where p is a reason for you to phi but where you falsely believe that p. But why not think that possession of a reason merely requires truly believing it?

---

<sup>6</sup> Another illustration of the flexibility of 'my reason': Suppose various people each have a reason for a certain bridge to be built pinned to their chest. Someone can in this context say 'My reason for the bridge being built is better than yours'. (Thanks to Tim Williamson here.)

<sup>7</sup> It is also worth keeping in mind that the possessive construction 'his reason for phi-ing was that p' is typically used to mark motivating rather than possessed normative reasons.

<sup>8</sup> We are attracted to a stronger claim still: (S knows P and P is a reason for S to phi) iff P is a reason to phi which S possesses. But we won't undertake to defend that here.

Suppose there is a dog in the room, which is standing behind a dog-shaped stone. Looking at the stone, a person X, who is afraid of dogs and mistakenly believes that the stone is a dog, runs out of the room. Upon which he claims: 'My reason for leaving the room was that there was a dog there'. To our ear, it is clear that X's statement cannot be true. But if true belief was sufficient for possession, then (since X does have true belief in this case), it is unclear why this proposition isn't available as X's motivating reason. On the other hand, if possession requires knowledge, then X does not possess the fact that there is a dog in the room as a reason, and hence this fact is not available as a motivating reason for X. Since availability as motivating reason is our key diagnostic for possession, we conclude that true belief is insufficient for possession. Exactly the same case, shows that even justified true belief is not sufficient for possession (after all, if the rock is a convincing facsimile the belief that there is a dog in the room will be justified as well as being true).<sup>9</sup>

As a way of blocking the dog-shaped stone argument, someone might appeal to the idea that a necessary condition on P's being X's reason for phi-ing is that p be one of the reasons *why* X phi-ed (i.e. motivating reasons have to be explanatory reasons).<sup>10</sup> In the case described, it is false that one of the reasons why the subject runs is that there is a dog in the room. It may be urged that it is this fact and not the absence of knowledge that makes for the falsity of the motivating reason ascription. But it is easy to describe somewhat similar cases where the purported motivating reason is an explanatory reason, but where, owing to the lack of knowledge, it is odd to describe it as a motivating reason. Suppose for example, that the dog had dragged the dog-shaped stone into the room, or that one suffers from a rare neurological condition which induces one to hallucinate dogs in the presence of dog-pheromones.

This concludes the initial case in favour of the claim that possession requires knowledge. In the following two subsection we address two objections to this claim.<sup>11</sup>

---

<sup>9</sup> This line of argument follows that of Unger (1978), Hyman (1999), and Hornsby (2007).

<sup>10</sup> A putative counterexample: A sexist employer promotes a man over a female competitor. Let us suppose in this case that the employer would cite as their own motivating reason the fact that the male employee gave an excellent speech at the annual general meeting. But suppose that the employer would, due to their sexist biases, have chosen the employee at all nearby worlds. It might reasonably be claimed that it is their bias and not the speech that was the reason why male was chosen for promotion. There are various ways to resist this attempt to block the entailment from motivating reasons to reasons why. Depending on how the details are fleshed out, this might be (i) a case where both the sexist bias and the speech are reasons why or (ii) as case where the employer is under an illusion, mistakenly thinking the fact about the speech was a motivating reason when in fact it was not.

<sup>11</sup> One objection that we will not consider at length is that possession merely requires being in a position to know (as opposed to actually knowing). One way of motivating this thought is by appeal to examples using the possessive construction as applied to normative reasons where it is obvious that the person does not know. (E.g. if there is a post-it note bearing bad news that is there to be read but not yet read, it doesn't sound too bad to say 'The agent has reason to be depressed'.) Such examples trade on the flexibility of the possessive construction noted earlier – we think it unpromising to look for a single epistemic relation that will accommodate both these examples and paradigmatic examples of possessed reasons, especially as the possessive-normative construction extends in certain contexts to cases where there is no interesting epistemic relation, potential or actual to the relevant fact. Another way to try and resist the claim that possession requires knowledge is to maintain that certain stative factives – such as seeing that – suffice to enable p to be possessed as a reason but don't require

## **§2.2 Motivating reasons in 'fake-barn' cases**

The first objection we anticipate rests on the claim that the knowledge requirement too demanding for a range of fake-barn style cases. Suppose you are looking at a tiger in fake-tiger country. You are, understandably enough, afraid of tigers and run. The report 'My reason for running was that there was a tiger in front me' sounds acceptable enough. But the majority of philosophers will judge that in such a case one does not know that there is a tiger there (since this is a situation where there are numerous fake tigers in the vicinity).

This style of argument can be generalised to counter claims of knowledge entailment for a large range of verbs. For even in such a 'fake-tiger county' setting, 'I saw that there was a tiger there', 'I realized that there was a tiger there', 'I discovered that there was tiger there', 'I was dismayed that there was a tiger there' all sound reasonably acceptable.

There are three reactions one might have to such cases. First one might think that, after all, one does know that there is a tiger in such a scenario (despite the fake-tigers in the vicinity). Second, one might think that the initial judgment of acceptability of the reason (/seeing/discovering/etc.) ascriptions is, while prima facie attractive, nevertheless incorrect. Third, one might think that these kinds of cases show that 'realize that', 'discover that', 'see that', and 'dismayed that' all fail to entail knowledge. We have some sympathy for the first option though we shall not press that view here. But at any rate it seems that the third reaction is the least plausible. It is very hard to accept that statements such as the following can be true: 'John discovered that there was a tiger there but never knew that there was a tiger there' or 'John realized that there was tiger there but never knew that there was tiger there'. It seems much more likely that the passability of 'I discovered that there was a tiger there' is traceable to the fact that the verb 'discover' draws attention to novelty and does not focus the knowledge question. (Notice that 'I knew that the tiger wasn't going to poison me' sounds much more passable than 'I knew that it was a tiger'. This is again because of facts concerning the focus of the statement.) The second response thus seems like a much more plausible way of explaining such judgements, and the objection fails to draw a wedge between possession and knowledge.

## **§2.3 reasons ascriptions in cases of false belief and second-order attitudes**

Let us turn to a second objection to the thesis that possession requires knowledge. In cases where someone falsely believes that p and where this belief is causally relevant to some further action, belief, or emotion, we tend to fall back on reason constructions involving psychological ascriptions as the complement of 'His/her reason that'. For example, if Jane falsely believes that there is a tiger in front of her and runs away, we will naturally resort to such explanations as 'Her reason for running away was that it looked as if there was a tiger in front of her' or 'Her reason for running away was that she thought that there was a tiger in front of her.'<sup>12</sup> (Such explanations satisfy the factivity demand of the

---

knowledge, because they don't require belief. Our view is that at best, the relevant considerations indicate that knowledge does not require belief. But pursuing this would take us too far afield.

<sup>12</sup> Notice that in some cases the 'looks' or 'appears' fallbacks are not a distinct option from the belief fallback. Consider a case where someone has a belief in preservative memory that fails to be knowledge and which is not

'reason that' construction since the psychological claims following 'reason that' are true.) The worry is that the truth of such ascriptions does not seem to depend in any way on whether the agent has a second-order attitude to the relevant psychological states, and in particular does not depend on whether the agent knows how things look or what she believes. The objection thus consists of two claims: first, that in such cases the agent's motivating reason is a psychological fact (Jane's reason was that she believed there was a tiger there), and second, that one can have such psychological motivating reasons without knowing that one is in the relevant psychological state.

Our main response to this objection consists of rejecting the first claim. We maintain that for a wide-range of cases where we are tempted to use ascriptions of the form 'Her reason for phi-ing is that p' with true psychological complements, the 'that'-clause does not ultimately succeed in picking out a motivating reason (and thus need not pick a possessed reason either). This means that even if such psychological facts are not known, this would not threaten the thesis that possession requires knowledge.

To begin, notice that when constructions of the form 'Her reason for believing P was that Q' pick out motivating reasons, they typically cite a proposition that is treated as evidence by the agent. But it doesn't seem (except in special cases) that the proposition that the agent believes that there is a tiger is treated as evidence by the agent.<sup>13</sup> This again suggests that such reasons ascriptions do not after all correctly ascribe motivating reasons.

One might think that 'It looked as if P' or 'It appears that P' are more promising psychological fall-backs in these cases, since they at least are better suited for an evidential role. Now certainly there are some cases where 'It looks as if P' is unproblematic as a motivating reason description. Suppose someone is unsure whether they are undergoing veridical perception of a tiger and thinking to herself "Well, at any rate, it looks as if there is a tiger there. So I'd better run just to be on the safe side." Here we can all agree that it is correct to cite the proposition that it looks as if there is a tiger there as a motivating reason. But things are far less clear in a case where the person takes the appearances at face value. To get a rough idea of how one might be sceptical here, consider an admittedly crude model of motivating reasons: People have an Aristotelian syllogism box in which various propositions appear as premises, instantiating practical syllogisms that yield action. A necessary condition for a proposition's being a motivating reason is that it figures as premise of practical reasoning in this way. In a case where the person takes the appearances at face value, the proposition that there is a tiger plausibly figures as a premise, but cannot express a reason because there is no tiger (at best that proposition is treated as if it were a reason). On the other hand, on this model, the proposition that it looks as if there is a tiger cannot figure as a motivating reason if it doesn't figure as a premise of

---

accompanied by any phenomenology. Here there is no belief-sustaining seeming fact distinct from the belief itself (i.e. no seeming fact other than facts such as 'it appeared to her that p') which can be appealed to as the person's reason for acting.

<sup>13</sup> As Hyman (1999) notes, one such special case is where X's reason for visiting a psychiatrist is that X believes she is being constantly followed.

sylogistic reasoning. And arguably, when appearances are taken at face value, the 'looks' proposition does not so figure in the agent's reasoning.

It is important to recognise, though, that even in those cases where such psychological facts do not constitute true motivating reasons, the ascription sentences (e.g. 'X's reason for running was that it looked as if P') might well have a true reading - it's just that on the true reading, the claim will not express a motivating reason ascription. Note that we routinely say such things as 'His reason for moving towards the fire was that he wanted to be warm' or 'Her reason for leaving was that she was bored' even in a case where the fact that the agent wanted to get warm/was bored did not figure as a premise in any kind of reasoning. Very roughly, the point of the possessive is to mark a psychological fact pertaining to the individual that is explanatory (i.e. a psychological reason why).<sup>14</sup> If there is indeed a mundane psychological-explanatory use of the possessive then there will be two very different jobs that might be performed by a claim like 'Her reason for phi-ing was that it looked as if P', one being to mark a motivating reason (which requires that the looks-fact is used as a premise in the agent's reasoning), another being merely to mark a psychological-explanatory fact (and hence where there is no requirement that the fact be known). To get a sense that such reasons-ascriptions indeed have two readings, contrast the following two cases in which Jill witnesses a convincing hologram of a tiger in the room. In the first case, she reasons to herself by saying: 'There is tiger in the room, so I should run'. In the second case, she reasons to herself by saying: 'I'm not entirely sure if there is or isn't a tiger in the room. But it looks like there is a tiger in the room, so I should run.' There is one reading of 'Jill's reason for running was that it looked as if there was a tiger there' on which the ascription is true of both cases (this is the psychological explanation-why reading), but there is also another reading of the sentence on which it would only be true in the second case (this is the motivating reasons reading).<sup>15</sup>

The upshot is that in a wide-range of cases, even when we are tempted to use 'thinks that' and 'looks as if' as fall-backs for motivating reason constructions (and even if these constructions have *some* true reading), it may not be correct to think that the complements of the that-clause are motivating reasons: where such psychological facts do not in fact figure in the agent's practical reasoning process, they cannot be motivating reasons.

---

<sup>14</sup> Matters are actually a bit more delicate. We tend to use such 'his reason for phi-ing' constructions only to explain actions are normatively evaluable. Thus there is an interesting contrast between, e.g., physical pain and emotional grief: 'Her reason for being in pain is that someone stepped on her foot' is, in most cases infelicitous, while 'Her reason for being in pain was that her father died' is entirely felicitous. Precisely what accounts for this contrast is a matter for a separate discussion.

<sup>15</sup> Thanks for discussions with Christina Dietz here. Hornsby (2007) argues that we cannot understand such ascriptions as expressing explanatory reasons because when we truly make such ascriptions as 'X's reason for running was that she believed there was a tiger there', those ascriptions are true in virtue of some reasoning process that X undergoes (this is contrasted with some other cases of explanation: if the reason the bridge collapsed was that it had a weak link, this is not true in virtue of any reasoning-like process on behalf of the bridge). But even if the psychological reasons ascription is true in virtue of some reasoning-like process on behalf of X, this does nothing to rule out that the reason is explanatory (explanations of different facts can be highly diverse, and in this case the particular explanation might appeal to X's reasoning).

Suppose we are in one of those cases where it is wrong to think that the proposition that it looked as if there was a tiger was a motivating reason in some hallucination case. What about a possessed normative reason? It does not follow that the looks proposition was not a possessed normative reason to run. For suppose that the proposition was known but not in fact motivating. Still, the agent *could* have used it as a motivating reason – for example in a situation where they were being cautious about whether there was in fact a tiger. In short, in a situation where a tiger is hallucinated (and the agent knows the looks fact), we have little doubt that the agent possesses a reason to run. Nevertheless, since in typical cases the ‘looks’-fact is not a motivating reason for the agent’s action, there is little pressure to count this fact as a possessed reason when the agent does not know this fact. This kind of case thus poses no threat to the claim that possession requires knowledge.

The objector might try to press the point, and maintain that even in those cases where such psychological reasons ascriptions do legitimately report motivating reasons, they do not require the agent to know the psychological complement. Thus, for example, the objector might insist that even in cases where Jill’s reasoning does involve premises such as ‘It looks as if there is a tiger there’, or ‘I believe there is a tiger there’, we need not require that Jill *knows* that these psychological premises hold in order for these claims to count as Jill’s motivating reasons.

Put this way, the objector’s worry seems far less compelling. Indeed, the same considerations that justified the claim that possession requires knowledge in the case of worldly facts, generalises to the case where the purported reasons are psychological facts. Consider for example the following case: Jill is an art dealer who is interested in buying work containing patches which look to be a particular fine-grained shade of red – red<sub>27</sub> (she has read of recent research showing that buyers are more likely to buy artwork that looks to have that particular shade, perhaps because they are causally sensitive in a certain way to its appearance). She comes across an artwork that in fact has a patch which looks the relevant shade, but despite seeing the patch, she does not know that it looks red<sub>27</sub> (suppose she cannot discriminate between cases where things look to be red<sub>27</sub> and cases where they look to be of a very similar shade such as red<sub>28</sub> or red<sub>26</sub>). Still, Jill might form a true belief - if you want, a justified true belief- that the artwork contains a patch that looks red<sub>27</sub> (suppose that under the picture is a label, claiming the patch on the left looks red<sub>27</sub>, but unbeknownst to Jill this label was accidentally left by the curator from a previous exhibition). Jill might use her belief that the patch looks red<sub>27</sub> in her practical reasoning and end up buying the picture. But still it would be false in this case to say that her reason for buying the picture was that the patch looked red<sub>27</sub>.<sup>16</sup> Finally, note that the objector’s line here (which maintains that only when psychological facts are ascribed as reasons, knowledge is not required for possession) risks making the possession relation highly disjunctive. After all, as we have argued, in order for an external fact (or even an internal physiological fact) to count as a possessed reason, one has to stand in the knowing relation to it (cf. the dog-shaped stone case above). Yet the

---

<sup>16</sup> Of course, if in addition to her Gettiered-belief about the patch’s colour, Jill is also causally sensitive to the patch’s colour, the sentence would have a true reading as a psychological-reason-why ascription. (But note that even in that case, there will also be a second reading – the motivating reasons one – on which the ascription would be false.)

objector's position seems to be that for a belief-fact or looks-fact one needn't stand in the knowledge relation to that fact in order for the fact to count as a possessed reason. But this would seem to require that possession is a gerrymandered relation, which would certainly be a cost of that view.<sup>17</sup>

We have argued that in typical cases where an agent hallucinates a tiger, reasons 'There is a tiger there so I had better run', and then runs away the agent did not have the 'looks as if' or 'believes that' as motivating reasons. Moreover, given our commitment to the factivity of reasons (see §3 below), it does not seem that in such cases there are any *other* suitable candidates for being the agent's motivating reasons. Thus one upshot of our argument is that typically, in such cases the agent acts for *no motivating reasons*. An obvious worry about this proposal is that in many such cases we might be tempted to say that the agent is acting rationally, and one might worry that one cannot act rationally if one acts for no reason. We will return to this objection at the end of §3 below, so will not discuss it any further here.<sup>18</sup>

In summary, we maintain that reasons ascriptions with psychological complements such as 'His reason for running was that he believed/it looked like there was a tiger there' pose no threat to the thesis that possession requires knowledge: either (as is the more typical case) they do not pick out a motivating reason of the agent and thus the thesis does not predict that they require knowledge, or they do correctly describe a motivating reason (as might be the case in situations where the agent is being cautious and uses the psychological fact in their practical reasoning) – but then the agent really is required to know the relevant psychological fact.

### §3. Factivity and Factoring

The case for the factivity of 'A reason to phi is that P', 'His reason for phi-ing is/was that P', and 'A reason he has to phi is that P' seems just as compelling as for paradigmatically factive attitudes such as 'knows that' and 'regrets that'. Sentences like 'His reason for leaving was that the house was on fire but the house wasn't on fire' sound terrible in such the way that 'He regrets that the house was on fire but it wasn't' sound terrible. Meanwhile, the inference from 'A reason he has to phi is that P' to 'P' strikes us as deductively valid in such the way that the inference from 'He regrets that P' to 'P'

---

<sup>17</sup> A similar worry is expressed by Schroeder (2008) (Schroeder does not himself endorse the thesis that possession requires knowledge, but he does maintain that in the case of worldly-facts, to have p as a reason one should stand in some epistemic relation to the reason - e.g. believing p, and he maintains that if a psychological fact is to act as a possessed reason, the agent would be required stand in the same epistemic relation to the relevant psychological fact.)

<sup>18</sup> The claim that there can be cases where an agent falsely believes that they are acting for a motivating reason might lead the reader to suspect that we are defending here the view that all motivating reasons are normative reasons. This, however, would be a misunderstanding of our argument: what we have argued for is merely the claim that motivating reasons have to be both known and in practice play the right kind of motivating role in the agent's action. This still leaves open the view that a known fact can be a motivating reason to phi even if it does not count in favour of phi-ing, and hence is not a normative reason. For example, if the reason for which a sexist employer chooses not to promote a female employee is that the employee is a woman that might be a motivating reason that is not a normative reason.

strikes us as deductively valid.<sup>19</sup> Of course there are true generic readings of sentences such as ‘A reason to take a teaching job is that it offers flexible hours’, and one can truly utter such a sentence even on an occasion where one is contemplating a particular teaching job that does not offer flexible hours. But that does not make trouble for factivity any more than the fact that one can truly utter the true generic ‘People remember that they went to school’ despite the fact that there are certain people who never go to school.

It is worth noting that there is an acceptable use of ‘His proof that P’, ‘His memory was that P’, and ‘His explanation was that P’ which do not imply that P is true or provable. (Think of ‘His proof that God doesn’t exist was terrible’ and ‘His memory is he that he never made that promise’.) The use of the possessive in these cases is to signal, roughly, that something *seemed* like a genuine proof or a genuine remembering from the subject’s point of view.<sup>20</sup> Of course this is not the only use for the possessive in connection with ‘proof’. There is a standard reading on which ‘Andrew Wiles now has a proof of Fermat’s last Theorem even though he hasn’t proven that it is true’ or ‘Her proof of the Pythagorean Theorem is shorter than his, but neither have proven the theorem’ sound entirely defective. Perhaps the minority of philosophers who think that the possessive reason construction is non-factive are hearing a use of ‘Her reason’ that works along the perspectival lines indicated above. Note, though, that as with ‘proof’, even if there is such a use of the possessive, this is certainly not the only use – indeed the factive use seems to be more characteristic and the one of greater theoretical interest. Also, even if there is such a non-factive use, this would not mark an ambiguity distinctive of ‘reason’. As we have seen above, we can, for a wide variety of nouns N, use ‘Her/His N’ to characterize something that seems to be N from the subject’s perspective. This does not point to a theoretically interesting ambiguity in ‘reason’ any more than it does so for ‘proof’ and ‘explanation’.<sup>21,22</sup>

Granted that various ‘reason’ constructions have at least an air of factivity, one might try to explain this away by maintaining that while ‘reason that P’ does not semantically entail that P, such reason ascriptions somehow pragmatically communicate that P. It seems highly dubious that this so called

---

<sup>19</sup> It goes without saying that in a case where an agent is mistakenly takes themselves to know that P even though not-P, they will happily report themselves by such constructions as ‘My reason for phi-ing was that P’. This is not more evidence for the non-factivity of this reason construction than the fact that such an agent might say ‘I know that P’ is evidence that knowledge isn’t factive. (We mention this partially because Schroeder seems to be misled into thinking that such perspectival reports provide some encouragement for the claim that possessed reason ascriptions are non-factive (see Schroeder (2008), p.10).)

<sup>20</sup> Cf. Hyman (1999), p. 445.

<sup>21</sup> Also relevant are the following uses of definites and demonstratives: ‘The/that diamond is fake’, ‘The/that proof is terrible’. A systematic treatment of the nature of proofs should not give such uses central stage any more than a systematic treatment of the nature of diamonds ought to.

<sup>22</sup> One of our main opponents is Mark Schroeder (Schroeder (2008)), who argues that propositions don’t have to be true in order to be what he calls ‘subjective’ reasons and where there is a fundamental ambiguity in ‘reason’ between the factive ‘objective’ use, and the non-factive ‘subjective’ use of ‘reason’. Even if one granted a non-factive use of ‘His reason’ of the sort suggested above, that would not do much to advance Schroeder’s vision. For first, that use does not point to an interesting ambiguity of ‘reason’. Second, pace Schroeder’s discussion, it does not suggest that possessive reason constructions are in general non-factive. And third, Schroeder seems to suggest that the use of ‘His reason’ he is interested in does at least carry an *implication* of factivity (even if it does not strictly entail factivity). On the other hand, the perspectival use of ‘His reason’ discussed above carries no such implication.

implication is a mere conversational implicature. (It is not clear what general conversational maxims would be flouted, and moreover the supposed implicatures do not appear to be cancellable.) Perhaps one might claim that the sentence lexically triggers a presupposition of factivity. But note first, that typically, the presuppositions of a lexical item tend to also be semantic entailments in simple non-embedded declarative environments.<sup>23</sup> So far example, while it is widely accepted that 'X knows that P' and 'X regrets P' presuppose that P, it is also clear that they also semantically entail that P. Admittedly there is one kind of lexically triggered non-semantic implication that is not accompanied by entailment even in unembedded environments, namely conventional implicatures. (Thus, for example, 'He is poor but happy' is usually taken to conventionally implicate but not semantically entail that there is some tension between being poor and happy.) But the move which maintains that reasons constructions carry a conventional implicature of factivity seems entirely ad-hoc unless someone produces some data that would encourage the view that while 'X knows that P' entails P, 'X's reason was that P' conventionally implicates but does not entail P.<sup>24</sup>

A related question discussed by Schroeder is whether possessive reason constructions can be factored into a conjunction of a reason claim and a possession claim.<sup>25</sup> For example, does 'She has a reason for leaving, namely that there was a tiger in the room' factor into 'There was a reason for her to leave the room, namely that there was a tiger, and she was in possession of that reason'? Putting aside the alleged perspectival reading noted above, we maintain, pace Schroeder, that possessive reasons do factor in this way.<sup>26</sup> Certainly the superficial linguistic data strongly supports this view. It sounds very strange to say 'He had a reason to leave though there was no reason for him to leave'. This strongly suggests that 'He had a reason to leave' entails 'There was a reason for him to leave'. Indeed the factorization structure does not seem in any way peculiar to reasons talk. In general, 'X had an F' entails there is an F and X had it. (e.g. 'He had a valuable ring' entails 'There was a valuable ring and he had it'.)

It is important not to conflate the factorization question with another question, namely whether the status of P as a reason is independent of P being possessed. Consider, by analogy, the property of being valuable. Some objects may count as valuable independently of who they are possessed by. But suppose for example that any object possessed by Paul McCartney is ipso facto valuable. Consider a

---

<sup>23</sup> See Abbott (2006) for the claim that this is true of all presuppositions.

<sup>24</sup> Moreover the factivity implication for reasons ascriptions fails the standard tests for being a conventional implicature (see Potts (2007)). Here is one such test: conventional implicatures (but not entailments or presuppositions) seem to survive under propositional attitude reports. Thus for example with implicatures generated by appositives we get: 'John believes that Lance Armstrong, that cheater, deserves his medals' commits the speaker to the claim that Lance is a cheater. Note that the factivity of reason-constructions fails this test: 'John thought that Mary's reason for leaving was that she was bored' does not seem commit the speaker to the claim that Mary was bored.

<sup>25</sup> See Schroeder (2008), who argues that they cannot be so factored.

<sup>26</sup> On any use of 'There was a reason, namely P' that is factive, the perspectival 'His reason was that P' will not factorize, since the latter but not the former can be acceptable even where not-P. (Note that this feature is shared with the perspectival readings of other constructions – on the perspectival reading, 'His proof of p was long' doesn't entail 'There was a proof of p', so cannot be factorised into a conjunction with this as one of the conjuncts.)

biro whose value consists entirely in its being possessed by McCartney. Its status as valuable is not independent of their possession. Had Paul McCartney not possessed it, it would not have been valuable. None of this conflicts with the factorization of 'McCartney has a valuable biro' and no ambiguity needs to be postulated for 'valuable' to account for the relevant facts.

Similarly, we think that the status of some propositions as reasons depends on their being possessed. Suppose we in a context where the guiding normative ordering source is subjective in the sense explained in §1 above. In that context, a fact will be a reason only if it bears on what is what is most likely to achieve the relevant ends given the subject's evidence. Suppose that in such a context, a subject knows that there is a tiger in the room and runs. In that context 'The fact that there was a tiger in the room was a reason to run' is true, but it would not have been true if the agent had been oblivious to the presence of a tiger. For holding the subjective ordering source fixed, the sentence 'That there is a tiger in the room is a reason to run' is false at such worlds. (Of course such a sentence might be true relative to a more objective ordering source.)<sup>27</sup> In this case, then, the status of the proposition that there is a tiger in the room as a reason for the agent to run is dependent on that proposition's being known and hence possessed.

In attempting make trouble for factorization, Schroeder suggests an analogy between possessive reason talk and certain other possessive noun-phrase constructions such as 'His father' and 'Her golf partner'. One's status as a father or as a golf partner does depend on there being someone that one is a father or golf partner of. Note that this does not strictly speaking conflict with factorization per se. After all, it is true enough that if X is Y's father, then X is *a* father. Moreover we think 'reasons' talk is more akin to 'valuable' than to 'father'. No father can be a father without being the father of someone. But many reasons can be reasons without being possessed by someone. This is obvious when the ordering source is a more objective one. (That something is poisonous is a reason for someone not to drink it even if that reason is not possessed.) Even for a subjective ordering source, not all reasons need to be possessed. Consider 'That P is unlikely on his evidence is a reason for him not to be very confident about it'. This can be true relative to a subjective ordering-source even when the agent in question does not know that P is unlikely on his evidence.

Schroeder's chief argument against the factorization view concerns a famous case (originally due to Williams) where someone is given a glass which in fact contains gasoline, but which the agent mistakenly believes contains gin and tonic and because of this drinks the liquid. Schroeder's argument seems to proceed as follows. First, given that the agent was not being irrational in this case, she must have acted for a reason (i.e. there was a motivating reason for the agent's action), and moreover the reason for which she acted must be a (possessed) normative reason. But then if we assume factivity (which Schroeder presents as driven by factorization) for possessed normative reasons, this reason must be a true proposition and, Schroeder argues, no such proposition is a good candidate for being

---

<sup>27</sup> Note that we don't wish to claim that for a subjective ordering source, the only way that some P can function as a reason is by being known. We return to this point below and in §4.

the agent's reason for drinking in this case. (The main candidate he considers is the proposition that the agent believes that the glass contains gin and tonic.)

Our main objection to Schroeder's line here is that we think that despite the temptation to say that the agent was not acting irrationally there need not be a motivating reason for their action (let alone a motivating reason which is also a normative reason). We will return to this point below, but before it is worth asking whether – rationality considerations aside – the agent indeed has any motivating or possessed normative reasons for drinking in this case.

With respect to motivating reasons, while we agree that there is a strong prima facie temptation to think that something was the agent's motivating reason for acting, this may be due to the fact that we are invariably content with such fallbacks as 'The agent's reason was that she believed that/it seems that the glass contained gin and tonic'. But as we noted above such fall-backs may not survive critical reflection. And if we are willing (as Schroeder himself does) to jettison these fall-backs and maintain that the 'believed that' or 'it seems that' claims weren't after all the agent's motivating reasons, we ought to take seriously the hypothesis that there was no reasons for which the agent acted (though of course there were reasons why she acted the way she did and though there may be propositions the agent treated or thought of as reasons).

What about possessed normative reasons? While we are inclined to accept that the agent did not act for such reasons as that the agent believed the glass contained gin and tonic or that it seemed to contain gin and tonic (i.e. these were not the agent's motivating reasons in this case, and arguably, there were simply no reasons for which the agent acted), we do think that certain of these propositions were reasons the agent had for acting. (By analogy, in the case discussed above where one hallucinates a tiger, that there looks to be a tiger is a reason for the agent to run even if it did not in fact serve as the agent's reason for running.) Moreover we find Schroeder's objections to counting such psychological propositions as possessed normative reasons as unpersuasive.

Focussing on the suggestion that the possessed reason is the fact that the agent believed the glass contained gin and tonic, Schroeder considers an enlightened bystander who is asked to tally the reasons for and against drinking the liquid. Given factorization, if the fact that the agent believed there was gin and tonic was a reason he had to drink it, then that fact was a reason for the agent to drink it. But it would be very strange, Schroeder notes, for the enlightened bystander to cite the fact that the liquid was gasoline as a reason for the agent not to drink the liquid, while at the same time citing the fact that the agent believed the liquid to be gin and tonic was a reason for the agent to drink it. This strangeness persists for such suggestions as that the fact that the liquid appeared to be gin and tonic, or that the agent had good evidence that the liquid was gin and tonic, were reasons for the agent to drink it. Thus if Schroeder's argument works, it refutes all these other suggestions as well. The problem with the argument, however, is that it abstracts away from fact that in addition to the question whether a reason is possessed or unpossessed, there is also the question whether the relevant normative ordering-source is subjective or objective. Assume first that we are reading 'reasons to phi' using a subjective ordering-source. In that setting, the fact that the liquid is gasoline has little bearing

on whether the agent 'ought' (read subjectively) to drink, and thus would not figure at all in the bystander's table of reasons for and against the agent's drinking (even though the bystander knows that the glass contains gasoline). Suppose instead that the bystander uses an objective ordering-source. Then arguably, the fact that the agent believes the glass contains gin and tonic is no reason at all for them to drink.<sup>28</sup> Thus, even with factorization and factivity, we can explain the relevant strangeness by suitable appeal to the candidate ordering-sources.

Here is a further worry about the idea that in the bad case where, e.g. a tiger is hallucinated, the proposition that it looks as if there is a tiger in front of a certain agent is a reason that agent has to run. The concern is that we are left with no happy view of what to say in the good case where the person knows there is a tiger. For first, it seems strange to say that the looks proposition is a reason the person has to run in the bad case but not in the good case. But second, it seems strange to say that in the good case the person has two reasons to run – (i) that there is a tiger there and (ii) that it looks as if there is a tiger there. In response, we agree that the proposition that it looks like there is a tiger is available in both the good case and the bad case. Indeed, as we noted, the cautious person may well use the looks proposition as their motivating reason even in the case where they know. (After all, they may be unsure whether they know...) So we do wish to claim that in the good case, both propositions are reasons the person has to run. Any sense of strangeness in this claim arises from a failure to notice one of two things. First, while each of the two propositions are available to the agent as suitable motivating reasons to run, we agree that it may be atypical that the person uses both propositions as motivating reasons.<sup>29</sup> Second, there are many cases where a person has two reasons for acting a certain way or for believing something but where one is so much more decisive than the other that it is pragmatically strange to cite both. That the liquid tastes very bitter and that it contains arsenic are both reasons not to drink it. But it would be somewhat odd to say 'The agent has two reasons to refrain from drinking that liquid – that it contains arsenic and that it tastes very bitter'. We conclude, then, that the most plausible thing to say about the gin and tonic case is that the agent did possess a normative reason to drink, but that reason did not serve as the agent's motivating reason.

We close this section by discussing what is arguably the most pressing challenge for the view we have been proposing: namely the purported connection between motivating reasons and rationality. Our

---

<sup>28</sup> This at least seems to be plausible if we adopt the view of reasons as explanations (cf. §4 below) – that the agent believes the glass contains gin and tonic arguably has no explanatory force in support of the claim that they ought (objectively) to drink. (The argument is somewhat complicated because it involves the controversial case of a pro-tanto reason. But here is a similar case that illustrates the point without appealing to pro tanto reasons: the glass in fact contains gasoline. Unbeknownst to the agent, drinking gasoline would make them feel very unwell but ultimately save their life from a rare disease. The agent falsely believes the glass contains gin and tonic, and thereby drinks. Here the agent ought (objectively) to drink, but the fact that they believe that the glass contains gin and tonic is in no way part of the explanation of why they ought to drink.)

At any rate, even if one wishes to count the belief-fact as a reason to drink, it at most would act as a very weak and trivially overridden reason. Either way, one would predict some discomfort in counting the belief fact as having any substantive force to the 'pro drinking' side.

<sup>29</sup> Unless of course one thought that (i) the knowledge that there is a tiger is inferred from knowledge that it looks as if there is a tiger and (ii) that this inferential structure is sufficient for both to count as motivating reasons.

suggestion was that that in the gasoline case (*mutatis mutandis*, in the typical case where one hallucinates a tiger and runs), the agent acts for no motivating reasons. But the natural worry here is that the agent in such cases does not seem to be irrational, and there is a strong temptation to accept that anyone who acts for no motivating reasons is irrational. Indeed, Schroeder pushes this line even further by arguing that the only way of rescuing the agent from a charge of irrationality, is to allow false propositions to function as both motivating and possessed reasons.

Even setting aside the fact that there is overwhelming evidence that our reasons talk is factive (see §3 above), it is not clear that allowing false propositions to function as motivating reasons provides a good theory for the connection between reasons and rationality. Indeed, such a view risks erring in the opposite direction – it will count people as rational when they are anything but. Think of people who have all sorts of belief pop into preservative memory out of nowhere on a regular basis and, because of this act in a very haphazard way. Or think of someone who has utterly perverse moral views which shape his pattern of behaviour. It is not particularly attractive to say “Such people are fully rational because they have excellent reasons for doing what they do”. But by allowing any false proposition in as a reason one risks licensing speeches like this.<sup>30</sup>

What is going on here we think is that there are at least two modes of evaluating agents which tug in different directions. For any property of an agent that we deem desirable we will first think it a good thing that an agent manifest that property but second, we will also deem it desirable that (roughly speaking) an agent acts in ways that would constitute manifesting that property under normal

---

<sup>30</sup> A common reaction we received to this line is that the problem can be avoided if we allow false propositions to count as motivating reasons only when the agent *justifiably* believes them. (Alternatively, one might be willing to count unjustified false propositions as a motivating reason but allow only the justified ones to be the motivating reasons of *rational* agents. Our remarks below apply equally to both proposals, as we are objecting to the claim that allowing justified false beliefs to count as motivating reasons provides a way of securing the connection between acting on otherwise acceptable reasons and the rationality of the agent.)

Discussing this suggestion in full would require more space than we have here, but here are a couple of considerations that tell against it. First, suppose that an agent knows P, P provides reasonable support for a false proposition Q which the agent then justifiably believes, and Q logically entails R. If we then allow Q to act as a motivating reason for the agent in question, then since Q logically entails R, Q is a completely decisive reason for the agent to believe R. But it is clear that, even if the agent is not being entirely unreasonable to believe R, she does not have a decisive reason to believe R either.

Second, our opponent faces a dilemma here. Either the notion of justification in play is very minimal, in which case it would leave the revised suggestion with exactly the same problem as the original one. (For example, if all that is required is internal coherence of one’s beliefs, then we can imagine an agent that has complete sets of internally coherent beliefs pop up out of nowhere and haphazardly into their preservative memory. Such an agent would hardly be acting rationally.) Alternatively, if the notion of justification in play is quite demanding, then even an agent that is intuitively (in one sense, at least) rational can often form unjustified beliefs. Thus for example, if P is taken to justify Q just in case P provides a level of (objective) evidential support for Q that is above a given threshold, then in cases where P provides evidential support for Q that is just barely below the threshold, a reasonable agent with good intentions could easily think that P justifies Q and come to believe Q. But then there will be many cases of intuitively rational agents that act on propositions that they are unjustified in believing, so we are back to a position that allows for ‘reason illusions’: sometimes agents who we would intuitively take to be rational nevertheless act for no motivating reason. This is a claim we ourselves accept, but it was precisely this consequence that the denial of the factivity of reasons was intended to avoid.

conditions. So, for example, a brain in a vat can get merit points on the second dimension even if not on the first with respect to the property of “being generous towards other people”.<sup>31</sup> This distinction carries over to acting for reasons: we take it to be a good thing for people to act for good reasons. But we also take it to be a good thing for people to act in ways that under normal circumstances would constitute acting for good reasons. The hallucination cases without motivating reasons that we are inclined to count as rational are ones where, roughly speaking, the agent scores well on the second mode of evaluation even if they score poorly on the first. In general there is a danger of an all-purpose term like “rational” being overworked. One risks lapsing into incoherence if one takes it as a necessary condition for rationality that one has good motivating reasons but sufficient for rationality that one does well by the second mode of evaluation. Perhaps it would be better to drop the term ‘rationality’ and more explicitly distinguish the various layers of evaluation. Allowing in false propositions as reasons is in effect a misguided attempt to collapse kinds of normative propriety that ought not to be conflated. We thus conclude that both factivity and factoring should be accepted.

#### **§4. Reasons: explanation or evidence?**

An interesting proposal defended in a series of papers by Stephen Kearns and Daniel Star is that a proposition is a reason for X to phi iff it is evidence that X ought to phi.<sup>32</sup> This proposal is developed in a way that respects factivity and factoring and can be rendered consistent with the distinction between possessed and unpossessed reasons, so long as one accepts a parallel distinction between possessed and unpossessed evidence.<sup>33</sup> We think that this proposal has quite a lot going for it – indeed we are quite unmoved by many of the concerns in the literature. Moreover, we think a position along these lines could be made stronger by recognizing context-sensitivity in the normative force of the infinitival ‘to phi’. (For example, there may be some contexts in which normative force focuses on what’s best all things considered, while others in which it corresponds to what is morally obligatory. This would take care of the worry that phi-ing can count as a reason for a supererogatory act, even though on one standard reading, it’s not evidence that one ought to perform it).

However, the view also faces some serious challenges. One worry has to do with how the view construes the notion of unpossessed normative reasons. As noted above, this requires us to accept some parallel notion of unpossessed evidence – evidence that is ‘out there’, independently of the epistemic situation of any particular agent. Some such notions are not foreign to the literature in epistemology. Thus for example, one might adopt Williamson’s hypothesis that there is an objective agent and world independent evidential probability function  $P$ , and maintain that a fact  $e$  is evidence

---

<sup>31</sup> Cf. Williamson on primary vs. secondary norms (Williamson (2000), ch.11).

<sup>32</sup> See especially Kearns & Star (2008), (2009), and (2013).

<sup>33</sup> Though see our concern about the latter below.

for a hypothesis  $h$  if  $P(h | e) > P(h)$ .<sup>34</sup> But note, first, that adopting Williamsonian evidential probability is far from an uncontroversial commitment (do we really want to accept that there is an objective, world-independent evidential probability that there are, say, pink elephants?...). More importantly, it is highly doubtful that this notion of evidence will suffice for Kearns and Star's purposes: suppose for example that a glass contains arsenic, but an agent  $X$  does not know this. That the glass contains arsenic is an (unpossessed normative) reason for  $X$  to refrain from drinking, and thus Kearns and Star maintain that the fact that the glass contains arsenic is evidence that  $X$  ought not to drink it. But there is no reason to suppose that a Williamsonian probability distribution would deliver this verdict (recall that, for example, in many possible worlds arsenic is extremely good for  $X$ 's health...). Rather, it seems that this fact is evidence only relative to some background body of information (which includes, among other things, the proposition that arsenic is poisonous).<sup>35</sup> That is not to say the view of reasons as evidence cannot be patched up with a theory of objective evidence (perhaps a contextually relevant background body of information might do some work here)<sup>36</sup> – but there is more to be said in developing this aspect of the view.

The second, more serious challenge for the view, concerns a certain kind of potential counterexample. Consider a case where a subject is holding an apple in their hand and wondering whether to eat it. Suppose that unbeknownst to the subject a highly reliable book hidden in a cave on the other side of the world says that the apple is poisonous (and also says, using an objective 'should', that no one should eat the apple.) If the apple is indeed poisonous we are happy to allow a context where it is true that the fact that the apple is poisonous is a reason for the agent not to eat the apple (using an objective 'ought'). But whether or not the apple is poisonous, it is hard to recover a context where it seems true to say 'That a reliable book said that the apple was poisonous was a reason for the agent not to eat the apple'. On the other hand it seems entirely natural to say that the fact that the reliable book says that the apple is poisonous is evidence that the agent ought not to eat it. (Similarly the fact that the reliable books says that the agent ought not to eat it is evidence that the agent ought not to eat it.)

Kearns and Star offer two responses to this kind of example. The first assumes that the force of the example relies on the fact that even a reliable book might be fallible. They then argue that all things considered this cannot knock out the candidate reason since all sorts of reasons for doing things are not in themselves conclusive. We agree that reasons do not have to be conclusive. Indeed, one way to see that focussing on the fallibility of the book is a red herring is to note that we can generate similar examples using factive verbs. To our ear, if a hermit in a cave on the other side of the world knows that the apple is poisonous, it remains hard to recover a context where it seems true to say 'That

---

<sup>34</sup> See Williamson (2000), chapter 9-10, though note that Williamson himself does not make use of the idea of unpossessed evidence. (Unpossessed evidence does not satisfy the functional roles that Williamson takes to be key to a proposition's being evidence, e.g. the role of justifying belief.)

<sup>35</sup> And obviously, it would not due to just look at  $X$ 's knowledge as the background body of evidence (after all,  $X$  has a reason to refrain from drinking, even if  $X$  doesn't know that arsenic is poisonous).

<sup>36</sup> Cf. Kratzer's notion of circumstantial modality.

someone knows that the apple is poisonous is a reason for the agent not to eat it' (though of course the fact about knowledge is evidence that the agent ought not to eat it).

In our view the force of the example does not turn on fallibility but rather on a different feature. Following John Broome and others, we are tempted by the thought that there is a deep connection between reasons and explanation.<sup>37</sup> The basic insight is that the fact that hermit knows that the apple is poisonous is unsuitable as part of the explanation of why (for the objective 'ought') the person ought not to eat the apple. By contrast, the fact that the apple is poisonous is just the sort of thing to appropriately figure in the explanation of why the agent ought not to eat the apple.

Kearns and Star offer a second response to the alleged counterexample, one that they think mitigates the force of explanation-based accounts of reasons. Their thought is that if the agent were to read the book in question, then the fact that the book says that the apple is poisonous would give them a reason to refrain from eating it. (What of the suggestion that the book provides the proposition that the apple is poisonous as a reason? Given factivity, that would not work in a setting where the apple is not poisonous. But in that setting there would still seem to be a sense that reading the book would provide the agent with a reason not to eat the apple.) We agree that in the case where the book is read, there is a natural reading of 'That the book says that the apple is poisonous is a reason for the agent not to eat the apple' on which it is true. However, in our view, this does not entail that in the case where the book is *not* read, the relevant fact about the book is a reason for the agent not to eat the apple.

Indeed, with a bit of care about the normative force of the infinitival 'to-phi' construction, the explanation based approach can predict all the data concerning this example very nicely. First, consider the case where the 'ought' is objective: in that case, the fact about the book is not a suitable part of the explanation and so does not count as a reason. Meanwhile, in the case where the 'ought' is subjective, it matters whether or not the subject knows that a reliable book says that the apple is poisonous. When this is not known, the fact about the book plays no explanatory role vis-à-vis the subjective 'ought', since, roughly speaking, it does not explain relevant features of the expected utilities of the actions available to the agent.<sup>38</sup> Meanwhile, when the fact about the book is known, it does seem that the fact can be explanatorily relevant.

One might object that it is not the fact that a reliable book says that the apple is poisonous, but rather the fact that the agent knows this fact, which is explanatorily relevant to the subjective ought-claim. However, we think in general that when an agent knows that *p*, the fact that *p* is explanatory of the

---

<sup>37</sup> See Broome (2004) and Broome (2013), as well as so-called 'right-maker' accounts such as that of Schroeter & Schroeter (2009)

<sup>38</sup> Though as we have noted above, it is not in general true that unknown facts cannot be explanatory in this way. For example, the fact that such and such is likely on the subject's evidence can, on our view, be explanatory even if the relevant probabilistic fact is not known to the agent. And the fact that an object is round might explain why it looks round even when it is not known to be round, where this looks fact may be in turn be important to the expected utility profile.

knowledge fact. (The causal theory of knowledge is driven by this insight, though we don't ourselves assume that explanations have to be of an efficient causal sort.) Note that there is nothing at all surprising about the fact that subjective 'oughts' can be explained by external world facts in the case where they are known. The same is true for completely mundane explanations of psychological attitudes and actions. For example, if an agent knows there is a diamond on the floor and is thereby excited, it is entirely natural to say that they are excited because there is a diamond on the floor or that they are bending down because there is a diamond on the floor. We conclude that the explanation account predicts that there is a natural true reading of 'There is a reason for the agent to eat the apple, namely that a reliable book says that it is poisonous' when the agent knows that a reliable books says so but no natural true reading when the agent does not know this.<sup>39</sup>

The explanation approach is thus a promising alternative to the evidence view, but it does face a serious challenge of its own. We cannot in general accept that p is a reason for X to phi just in case p explains (or partially explains) why X ought to phi, because p can be a reason for X to phi even in cases where it is not true that X ought to phi. Since the 'reason why' construction is factive in the sense that 'p is a reason why q' can only be true if q is true, the explanation approach – at least in its simplest form - fails. The evidence approach, on the other hand, is naturally suited to accommodate this structure (it is completely standard for p to be evidence for q even if q is not true). Broome is of course aware of this issue and offers the following response. Assume it is a fact that one ought to (/ought not to) phi. There is a particular kind of explanation of this fact: a 'weighing explanation'. Such explanations work by analogy to the process of placing weights on a balance: we place weights on either side of the balance, and if the weights on the left-hand pan add up to more than the weights on the right-hand pan, the scale tips to the left. By analogy, consider a certain kind of explanation for why one ought to phi: *"There are reason for you to phi and reasons for you not to phi. Each reason is associated with a number which represents its weight. The numbers associated with the reasons to phi add up to more than the numbers associated with the reasons not to phi. That is why you ought to phi."* (Broome (2004): 36-37). The idea then is that something counts as a pro tanto reason for X to phi, just in case it plays the "pro phi-ing" role in a weighing explanation, even if the explanation in question is ultimately explanation for one you ought *not* to phi.

The development of the explanation account in terms of weighing explanations has the disadvantage that it makes the view far less clean, but we think it faces a deeper worry: placing weights on a scale is simply the wrong model for how to think about reasons for-and-against phi-ing.<sup>40</sup> Consider some of

---

<sup>39</sup> One complication. Perhaps there is use of 'ought' that is intermediate between subjective and objective by being grounded not on the evidence the subject has but (roughly) on evidence that is easily available to the subject. (Think of a case where a note is on the fridge but not attended to that says 'The apple is poisonous'.) In this setting it might be natural to invoke the note in explaining what the agent ought to do.

<sup>40</sup> Broome acknowledges the limitations of the metaphor and hedges his view by saying: *"Such a strictly analogous explanation rarely seems appropriate. For one thing, it often seems inappropriate to associate a reason with anything so precise as a number to represent its weight. Secondly, although we can aggregate together the weights of several reasons, to aggregate them simply by adding up also often seems inappropriate. So-called 'organic' interactions between reasons often mean that their aggregate effect differs from the total of*

the central features of the mechanical case: placing weights on a balance is a monotonic process (in the sense that placing more weights on one pan can only increase the total weight on that pan), and it is additive (to determine how much total weight there is on one pan of the balance one simply adds the weights). Relatedly, the only feature that matters to the process of placing weights, is how much each weight weighs (if at some stage of the process one places a golden square weight of 5kg on the left-hand pan that would have exactly the same effect on the process as placing a silver round weight of 5kg on that pan).

On the other hand, the process of considering reasons for-and-against phi-ing does not respect these features. Suppose you are trying to decide whether or not to buy a certain item on e-bay. If the item is made of jello, this fact might on its own be a reason for you to buy it (you like eating jello). If the item is a car, this fact might on its own be a reason for you to buy it (you need a car to travel around town). But if we already take on board that it's made of jello, the fact that it's a car will no longer count as a reason for buying it (such car will be of little use in travelling around), and moreover this fact will undermine the claim that being made of jello is a reason for you to buy the item (the jello car would presumably be grubby from the road, so you would no longer want to eat it). Similarly, if you like red items, then the fact that an item is bright red might be a reason (of a certain strength) for you to buy it, and the fact that it's red might also be a reason (of a certain strength) for you to buy it. But the strengths of the two reasons do not add up (you don't have twice as reason to buy it because it's both red and bright red). Finally, considered on its own, the fact that an item is made of jello might be a reason of equal strength for you to buy it as the fact that it's manufactured by Raleigh (you have an equal need for jello items for your kitchen and for Raleigh items for your bicycle collection). But adding each of these two facts to my deliberation can have radically different effects on the ultimate process (it will matter a lot to your overall deliberation which one you take on board, if you later also consider the fact that the item is a bicycle...). One could object that the talk of a so-called deliberation process is an unhelpful metaphor here, but the same point can be made more precisely without talk of such a process: if pro tanto reasons worked by analogy to weights, one would expect them to obey certain principles. For example, one would expect them to obey the principle that if p is a reason to phi and q is a reason to phi then p and q is a reason to phi. But the above examples show that this principle fails (that the item is made of jello is a reason to buy it, that the item is a car is a reason to buy it, but that it's a car made of jello is not reason to buy it). We conclude that weighing analogy is simply the wrong way to think about pro tanto reasons.

The evidence view, by contrast, has exactly the right resources to deal with the above structure: evidence can be non-monotonic, non-additive, and the ultimate import of a certain piece of evidence depends on its specific content, not only on how strongly it supports the proposition in question (these features become clear when evidence is cashed out in probabilistic terms). But given our worry

---

*their weights.*" (Broome (2004), p. 37). Similar reservations are stated in Broome (2013), p. 52. What is not clear, though, is what survives of the weighing analogy after we seriously take these points on board.

concerning the evidence view (examples such as the one involving the book or hermit above), one wonders if there are more subtle ways to expand the explanation view so as to cover pro-tanto reasons, that go beyond the simple weights model. While we shall not undertake that project here we might mention a few bodies of work from which one might draw inspiration. First, causal forces display a phenomenon akin to pro tanto reasons. There can be a force inclining an object to  $\phi$ , even if it doesn't  $\phi$ . A simple weights model for causal forces will obviously be too naïve. But a view of causal forces as evidence does not seem particularly palatable either. (For example, it will overgenerate. An effect might be evidence of its cause, but that doesn't mean it exerted causal influence on its cause.) Perhaps some of the ideological structure crafted to describe how causal forces combine might be of help in the current context. (Note, though, that causal forces in physics can be represented by vectors which can be added in the usual way. On the other hand, it's not clear that causal forces at the macro level are amenable to a clear vector representation.) Second, there is a great deal of literature on the phenomenon of pro tanto reasons in the epistemological literature that deploys inference graphs that encode directed patterns and strength of support between "prima facie" reasons and various propositions as well as phenomena of defeat (nodes on inference graphs get marked as defeated or undefeated).<sup>41</sup> The focus in that context is of course on reasons to believe, but insofar as they prove illuminating in the original context of application, there is no obvious reason why structurally similar graphs might not be of use in the current context.<sup>42, 43</sup>

One additional aspect of the way we talk about reasons is that we sometimes talk about a reason being strong or weak, one reason being stronger than another, or a certain reason out-weighting another. It might seem that one advantage of the evidence view over the (suitably amended) explanation account is that it allows for a ready-made analysis of such talk of the relative strength of reasons (after all, evidence too comes in various strengths, one piece of evidence can out-weigh

---

<sup>41</sup> This kind of project is associated above all with John Pollock. (See for example his 1995). For critical discussion see Lasonen-Aarnio (2010).

<sup>42</sup> Perhaps one or other of these models might also be helpful in making out a distinction that we have not done much with, namely that between a proposition's being an overridden reason and its being no reason at all. Suppose that a house has an air conditioning unit. But suppose further that use of air conditioners of that model has been made illegal. In that setting it is natural to say that the fact that the house has an air conditioning unit is no reason at all to buy it (as opposed to its being some reason to buy it, albeit one that is overridden by other factors.) Thanks to Julien Dutant here.

<sup>43</sup> Another version of the reasons as explanation view that does not rely on the weighing model appears in Finlay (2014), chapter 4. Finlay's proposal is that a pro-tanto reason for  $s$  to  $\phi$  is not an explanation for why  $s$  ought to  $\phi$ , but rather an explanation of why it is good for  $x$  to  $\phi$ .

It is not clear, however, that this solves the problem: after all, if the medicine saves your life but is disgusting you have a pro-tanto reason to refrain from drinking it, but it is far from clear that it is good for you to refrain from drinking. Finlay's response to this worry is that something can be said to be 'good' in many degrees and relative to different ends – it is sufficient for a pro-tanto reason to be good "in some way and to some degree" (ibid. p.90). The problem however, is that this risks making pro-tanto reasons ascriptions far too weak (it is not clear that on this liberal interpretation we will get true readings of 'He has no reason to do this'). Relatedly, on this account it is not clear why considering one's various pro-tanto reasons should play any role in determining what one ought to do. After all, what one 'ought' to do depends on the specific ends determined by the contextual resolution of 'ought', and the fact that an action is good relative to *other* ends is irrelevant.

another, and so forth).<sup>44</sup> One should note however, that our talk of the relative strength of reasons for phi-ing is far too context-sensitive and multi-dimensional to map neatly to the strength of evidence that one ought to phi. To see some of the complexity, consider for example the following case. X is given a glass of liquid which is both tasty and such that drinking it will save X's life. Let us also suppose that there are no countervailing considerations against drinking. In this case, the conjunctive fact that the liquid is tasty and there are no countervailing considerations is completely conclusive evidence that X ought to drink. Similarly, the conjunctive fact that the liquid will save X's life and there are no countervailing considerations is also completely conclusive evidence that X ought to drink. Thus both facts provide equally good evidence (namely, conclusive evidence) for the claim that X ought to drink. But intuitively, the fact that the liquid will save X's life and there are no countervailing considerations provides a much better reason for X to drink than the corresponding fact about the liquid being tasty. Roughly put, what this suggests is that the relative strength of reasons depends not only on how well the reason fact supports the aim in question, but also on how much we value the aim. To see some further complications, consider the following example: suppose you enjoy fizzy drinks, and you enjoy pink drinks and you value both properties equally. Suppose also that there is no correlation between a drink being fizzy and it's being unhealthy, but that if the drink is pink it's slightly more likely to be unhealthy (though assume that there is no explanatory connection between the colour and the healthiness facts – it just turns out that manufacturers who produce less healthy drinks also tend to like the colour pink). In the absence of any other considerations, that a drink is fizzy provides stronger evidence that you ought to drink than that it's pink (because the fact that it's pink also slightly increases the probability that it's unhealthy – which is a consideration against drinking). But intuitively, given that you value both properties equally, that the drink is pink is an equally good reason for you to drink as that it's fizzy. Roughly put, what this suggests is that the relative strength of reasons depends, at least in some cases, more on the explanatory force of each reason rather than on its evidential force. Of course, these considerations are not in themselves objections to the evidence view per se (the view does not have to commit to analysing the strength of reasons in terms of the strength of evidence), but it does suggest that the evidence view does not have the advantage of offering a ready-made analysis of such talk.

Another worry that might arise at this point is that the two views are sufficiently vague that it is hard to see what's at stake here (this worry is especially pertinent barring a worked out theory of unpossessed reasons in the case of the evidence view, and of pro tanto reasons in the case of the explanation view). A helpful way to defuse this worry and get some traction in the debate, is to note

---

<sup>44</sup> Kearns and Star (2009), §2.4 make this argument. Of course, Broome's "weights model" seems particularly well-suited for the task of accounting for the strength of reasons, but as we explained above we think that this model fails for other reasons. It is also worth noting in this context that the fact that we sometimes talk of a reason being 'outweighed' by another should not be taken to lend too much support to the weights model. After all, we also talk of a piece of evidence outweighing another, but it's fairly clear that we should not opt for a weights model for evidence.

that there are a range of structural features that plausibly separate the explanation view of reasons from an evidence based one. Here are some such features:

- (1) Conjunctions with irrelevant facts: suppose that P is evidence that one ought to phi, and consider the conjunction of P with some completely irrelevant fact R (in probabilistic terms: assume that the conditional probability of the ought claim given P is the same as the conditional probability of the 'ought' claim given the conjunction). In this case, the conjunctive fact P AND R will also be evidence that one ought to phi. For example, if the fact that the apple is poisonous is evidence that one ought not to eat it, then the fact that the apple is poisonous and Paris is the capital of France is also evidence that one ought not to eat it.

Explanations, by contrast, do not tolerate the addition of irrelevant conjuncts: suppose that the fact that the apple is poisonous is part of the explanation of why one ought not to eat it, it is very odd to say that the fact that apple is poisonous and Paris is the capital of France is part of the explanation of why one ought not to eat the apple. The two views thus have conflicting verdicts on the question of whether the fact that the apple is poisonous and Paris is the capital of France is a reason for the agent to refrain from eating the apple.

- (2) Reflexivity of reasons: On many natural conceptions of evidence, any proposition (at least any non-trivial proposition), is evidence for itself. (In a probabilistic framework the key consideration here is that, for any proposition with a defined probability above zero, its conditional probability on itself is 1). The evidence based approach thus predicts that the fact that the agent ought not to eat the apple is a reason for the agent not to eat the apple.

By contrast, explanations are plausibly irreflexive in nearly all cases. Thus the explanation account predicts that in typical cases the fact that the agent ought not to eat the apple is not a reason for the agent not to eat the apple.

- (3) A reason for both ought-phi and ought-not-phi: On most conceptions of evidence, no proposition can be both evidence for p and evidence for not-p. This in itself does not rule out that a proposition can be evidence for both for the claim that one ought to phi and for the claim that one ought not to phi (if one had non-zero prior in the claim that one neither ought to phi nor ought not to phi, then learning the disjunction that one either ought to phi or ought not to phi can increase one's probability in both ought claims). However, assume we are in a setting where it is certain on one's evidence that either one ought to phi or one ought not to phi (and certain that it is not the case that both ought claims are true). In this setting, nothing can be evidence both for the claim that one ought to phi and for the claim that one ought not to phi and hence, on the

evidence view of reasons, nothing can serve in this situation as both a reason to phi and a reason not to phi.<sup>45</sup>

But this is not so on the explanation view. The argument depends, of course, on how we fill out the details of the account of pro tanto reasons, but plausibly any such way would allow that the same fact can, in the same context, have explanatory pull both towards the claim that one ought to phi and towards the claim that one ought not to phi. Thus for example, suppose in front of you is a glass containing some blood. Drinking blood is very disgusting but on this occasion it will save your life. The fact that the glass contains blood has explanatory pull both towards the claim that you ought not to drink (if there weren't countervailing considerations here, it would have explained why you ought not to drink...) as well as towards the claim that you ought to drink (after all, it partially explains why you ought to drink). The explanation view thus predicts that the fact is both a reason for you to drink and a reason for you not to drink. The evidence view by contrast predicts (fixing the assumptions above) that in this case that it is only a reason for you to drink.

For what it's worth, we think these differing predictions provide some encouragement for the explanation based approach over the evidence one, but we admit that the data in these cases is not entirely clear. The main purpose of these distinctions, however, is not so much to defend one side of the debate, but rather to show that vague as they are, there are substantive differences between the two approaches, and to point to the kinds of considerations that can be useful in advancing the debate.

In this section we have contrasted two main approaches to what it is for the fact that p to be a reason for X to phi: the evidence approach and the explanation approach. One might of course offer another account. Alternatively, one might argue that the concept of a reason for X to phi is a primitive concept that cannot be reduced to other, more basic ones.<sup>46</sup> While we are perfectly happy with the use of primitive concepts in philosophy, the highly gerrymandered nature of our reasons-talk should cause us to doubt that reasons are indeed such a basic concept.<sup>47</sup> At any rate, we hope our remarks in this paper have gone some way to bring some clarity to this disorderly area.

---

<sup>45</sup> This feature of the evidence view is also highlighted in John Brunero's piece in this volume.

<sup>46</sup> See e.g. Parfit (2011).

<sup>47</sup> Another reason to be initially suspicious about the thought that reasons-talk is fundamental is that it does not translate very neatly into various other languages. In Hebrew, for example, there is no very natural translation for sentences such as 'Her reason for running was that there was a tiger'. The closest translation would use instead of 'reason' a word which (at least on the face of it) literally means 'cause', and moreover drops the possessive altogether. No doubt, a lot more needs to be explored to flesh this into a full-fledged argument against reasons fundamentality, but we have some reason to worry about views that hold that English happens to carve reality at its fundamental joints, while other natural languages do not.

## References

- Abbott, B. (2006): 'Where have some of the presuppositions gone?', In B. Birner & G. Ward (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, Philadelphia: John Benjamins: 1-20.
- Broome, J. (2004): 'Reasons,' in J. Wallace, M. Smith, S. Scheffler, and P. Pettit (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, Oxford: Oxford University Press: 28-55.
- Broome, J. (2013): *Rationality Through Reasoning*, Oxford: Blackwell.
- Finlay, S. (2014): *Confusion of Tongues: A Theory of Normative Language*, New York: Oxford University Press.
- Hyman, J. (1999): 'How knowledge works', *Philosophical Quarterly* 49: 433-451.
- Hornsby, J. (2007): 'A disjunctive conception of acting for reasons', in A. Haddock and F. MacPherson (eds.), *Disjunctivism: Perception, Action, Knowledge*, Oxford: Oxford University Press.
- Kearns, S. and Star, D. (2008): 'Reasons: explanations or evidence?', *Ethics* 119: 31-66.
- Kearns, S. and Star, D. (2009): 'Reasons as evidence', *Oxford Studies in Metaethics* 4: 215-42.
- Kearns, S. and Star, D. (2013): 'Weighing reasons', *Journal of Moral Philosophy* 10: 70-86.
- Kratzer, A. (1981): 'The Notional Category of Modality', in H.J. Eikmeyer and H. Rieser (eds.), *Words, Worlds, and Contexts*, Berlin and New York, de Gruyter: 38-74.
- Lasonen-Aarnio, M. (2010) 'Is there a viable account of well-founded belief?', *Erkenntnis* 72: 205-231.
- Parfit, D. (2011): *On What Matters*, Oxford: Oxford University Press.
- Pollock, J (1995) *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press.
- Potts, C. (2007): 'Conventional implicatures: a distinguished class of meanings', in G. Ramchand and C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press: 475-501.

Schroeter, L. and Schroeter, F. (2009): 'Reasons as right makers', *Philosophical Explorations* 12: 279-296.

Schroeder, M. (2008): 'Having reasons', *Philosophical Studies* 139: 57-71.

Williamson, T. (2000): *Knowledge and its Limits*, Oxford: Oxford University Press.