

# A “Good” Explanation of Five Puzzles about Reasons

Stephen Finlay

This is the author’s final preprint version of a paper accepted for publication  
in *Philosophical Perspectives* 33, 2020.

ABSTRACT: This paper champions the view (REG) that the concept of a normative reason for an agent S to perform an action A is that of an explanation why it would be good (in some way, to some degree) for S to do A. REG has numerous virtues, but faces some significant challenges which prompt many philosophers to be skeptical that it can correctly account for all our reasons. I demonstrate how five different puzzles about normative reasons can be solved by attention to the concept of goodness, and in particular observing the ways in which it—and consequently, talk about reasons—is sensitive to context (ends and information). Rather than asking simply whether or not certain facts are reasons for S to do A, we need to explore the contexts in which it is and is not correct to describe a certain fact as “a reason” for S to do A. These five puzzles concern: (1) reasons for attitudes of the “right kind”, (2) evidence as reasons, (3) normative facts as reasons, (4) subjective reasons, and (5) attitudes as reasons.

What does it mean for something, like the fact that *rain is forecast*, to be a normative reason for an action, like *taking your umbrella*, or attitude, like *believing it will rain*? According to a popular view, concepts of “reasons” are all concepts of some kind of explanation.<sup>1</sup> But explanations of what? On one way of developing this idea, the concept of a normative reason for an agent S to do A is that of an explanation why it would be *good* (in some way, to some degree) for S to do A.<sup>2</sup> This *Reasons as Explanations of Goodness* hypothesis (REG) has numerous virtues and champions. But like every theory of normative reasons it faces some significant challenges, which prompt many more philosophers to be skeptical that it can account for all normative reasons. This paper demonstrates how five different puzzles about normative reasons can be solved by paying attention to how the concept of goodness is context-sensitive—and how this implies that talk about reasons is also context-sensitive. Rather than asking simply whether certain facts are reasons for S to do A, we need to explore how context affects whether it is correct to describe a fact as “a reason” for S to do A.

These five puzzles concern (1) *reasons for attitudes of the “right kind”*, which depend on the *fittingness* of an attitude to its object, but apparently not on the value of having that attitude; (2) *normative facts as reasons*; for example, the fact that doing A would be good, which is arguably a normative reason but apparently could not itself be an explanation why doing A would be good; (3) *evidence as reasons*: reasons that are evidence an action would be good, but apparently are not explanations why it would be

---

<sup>1</sup> Early proponents include Toulmin 1950, Davidson 1963, Williams 1979. I take it to be roughly equivalent to accounts of reasons as *answers to (‘why’) questions* (Hieronymi 2005), and *grounds/truth-makers* (e.g. Schroeter & Schroeter 2009); cf. Broome 2008, Kearns & Star 2008.

<sup>2</sup> See Searle 2001: 110, Gardner & Macklem 2004, Wedgwood 2015, Maguire 2016, and in my own work, Finlay 2001, 2006, 2012, 2014. Raz (1999: 23) and Dancy (2000: 7-8) say similar-sounding things but ultimately argue against REG; see especially Raz 2011, chs. 2, 11.

good; (4) *subjective reasons*, which may seem from the agent's point of view to be explanations why action would be good, but cannot genuinely be so because the action actually isn't good; and (5) *attitudes as reasons*: sometimes we identify an agent's reason to act as *that they believed that p*, when this psychological fact apparently couldn't be an explanation why the action would be good.

Intuitions about these kinds of reasons are contested, but this is a further datum I will attempt to explain—as a consequence of context-sensitivity—rather than ignore, explain away, or hide behind. While all five puzzles appear problematic for REG, each also presents a more general challenge to many other theories of reasons. In demonstrating how REG provides resources to solve them all, this paper is therefore not only a defense of REG against objections, but also one piece of a positive abductive case for REG over its competitors. I first (section 1) survey some virtues of REG, to justify readers' investment in the paper. Next (section 2) I tackle the first puzzle concerning reasons of the “right kind” for attitudes, showing how reflection on our thought and talk about goodness leads toward a theory of goodness as *end-relative* (ERT), and how this combination of theories, REG+ERT, yields a straightforward solution. In section 3, I explain how ERT implies that talk about goodness is also *information-relative*. The remaining sections (4-7) show how this information-relativity provides solutions to the four other puzzles in turn.

## 1. General Virtues of REG

What makes REG a promising theory, and justifies the effort to see whether its apparent problems can be resolved? Philosophers have been drawn to it for a considerable number of virtues, which I will briefly sketch from most general to most specific.

First, REG is an *informative thesis* which promises to explain what it is to be a reason, in contrast to the widespread view that this property or relation is an unexplainable primitive.<sup>3</sup> While it might be suspected of merely substituting one mysterious primitive (being-a-reason-for) with a different one (goodness or value), I'll propose a reductive analysis of goodness itself, yielding a fully reductive analysis of what it is to be a reason.

Second, REG is supported by *linguistic evidence*, together with other reasons-as-explanations theories. The same noun 'reason' clearly refers to explanations in its nonnormative (causal) uses, as in talk about reasons for the unusually hot summer, or reasons why people are prone to bias. Words can admittedly be ambiguous, but in addition to enjoying the advantage of simplicity over non-explanation theories, REG is supported by the robustness of this pattern of dual use across a range of languages—including languages in the Germanic family (German: 'grund'), Romance family (Spanish: 'razón'), and Semitic family (Hebrew: 'סיבה'/'siba').<sup>4</sup>

---

<sup>3</sup> E.g. Scanlon 1998, Skorupski 2010, Parfit 2011b, Raz 2011: 18.

<sup>4</sup> Like English, these languages have a distinct word for *explanation* allowing a contrast with *reason*: 'erklärung', 'explicación', 'hesber' (thanks to Julia Staffel, Jaime Castillo Gamboa, and Dorit Bar-On). This is potentially problematic, as it presumably marks an actual if subtle difference in meaning. To speculate, note that 'explanation' is ambiguous between (*explanation*<sub>1</sub>) an *act* type/token of explaining, and (*explanation*<sub>2</sub>) the *explanans* or *content* of a successful act of explaining. I suggest 'reason',

Third, REG (with other reasons-as-explanation theories) offers a natural explanation of the *holism* of reasons: i.e. that a fact that is a reason for doing A in one situation might not be a reason for doing A in a different situation.<sup>5</sup> This is because explanations are holistic in the same way. The fact that it is raining may be an explanation why it would be good to take your umbrella against default background information, but not against the information that gale-force winds are blowing. Similarly, it qualifies as a *reason* to take your umbrella against the former but not the latter information. Relatedly, it is debated whether a “reason” should be distinguished from background conditions, or whether strictly speaking a “reason” includes all such conditions.<sup>6</sup> This is parallel to a debate over whether an “explanation”, strictly speaking, contrasts with or incorporates such conditions. Reasons-as-explanations theories account for these shared features of reasons and explanations.

Fourth, REG is *extensionally promising*, our five puzzles aside. Paradigmatic examples of normative reasons for action are prime candidates for explanations of why acting would be good, in some way and to some degree. For example, the fact that it is raining is both a reason to take an umbrella, and an explanation why taking an umbrella would be good: it will prevent you from getting wet. (For economy I’ll omit the qualifier ‘in some way and to some degree’ below, although it is important since we can have reasons for actions that aren’t good simpliciter.)

Fifth, REG naturally explains why normative reasons (although not other kinds of reasons) have *weights*. Normative reasons ordinarily have a pro tanto character: there can be reasons both for and against an action, of different weights or strengths. Here REG has an advantage over theories of reasons in terms of *oughts*, for example.<sup>7</sup> It doesn’t follow from your having *a* (or “some”) reason to do A that you ought to do A, but only from the overall weight of all your reasons taken together (“most reason”) favoring doing A.<sup>8</sup> Goodness exhibits the same pro tanto character, as it comes in degrees. It doesn’t follow from A’s being good to some degree that one ought to do A, but this does arguably follow from A’s being *best*. REG easily explains reasons’ weight as a function of the degree of goodness they explain.<sup>9</sup>

---

like ‘grund’ and ‘razón’, has the meaning of *explanation*<sub>2</sub> but not of *explanation*<sub>1</sub>, making ‘explanation’ a dispreferred translation. Many languages have a dedicated term for normative reasons (e.g. Chinese: ‘liyóu’; compare ‘justification’), while some other languages don’t have a noun corresponding to (normative) ‘reason’, their closest equivalent being talk about *why* it would be good/ one ought to do A (Wedgwood 2015). John Broome reports (ms) that in the earliest recorded normative use of ‘reason’ in English from around 1225, the writer slides from ‘the reason why one ought to A’ to ‘the reason for A-ing’.

<sup>5</sup> E.g. Dancy 1992.

<sup>6</sup> E.g. Nagel 1970: 90-95, Crisp 2000, Raz 1999: 228, Dancy 2004, ch. 2, Raz 2009, Schroeder 2009a.

<sup>7</sup> For ‘ought’ analyses, see Toulmin 1950, ch. 11, Broome 2004: 34, Finlay 2001: 104. For discussion of this advantage of ‘good’ analyses, see Phillips 1987, Finlay 2006, 2014: 90f, Sinnott-Amstrong 2009, Brunero 2013, Maguire 2016. Broome offers a disjunctive ‘ought’ analysis in response; for objections see Schroeder 2007, Brunero 2013, Kearns & Star 2008, Raz 2011: 21f. However, REG faces a similar problem in reasons for actions that aren’t good to any positive degree. Ralph Wedgwood (conversation) responds that ‘good’ has a variable degree parameter; I’m unconvinced (see Finlay 2014: 41-2), preferring to understand these as explanations of goodness *ceteris paribus*, relative to a restricted information-base (see section 3)—a move admittedly also available to an ‘ought’ analysis. I accept that some reasons are verdictive rather than pro-tanto (Gert 2016), understanding these as a special case of explanations of goodness, i.e. of *bestness*, equivalent to *oughts* (Finlay 2014: 90-1).

<sup>8</sup> I won’t explore the mass noun constructions, “some/more/most *reason*” here; see Evers 2010, Finlay 2014: 91f, Wedgwood 2015, Fogal 2016.

<sup>9</sup> Evers 2010, Finlay 2014: 92, Wedgwood ms. Compare Maguire’s (2016) suggestion that reasons’ weight is a function of goodness *and* probability; cf. Raz 2011: 115. The end-relational analysis sketched below incorporates probabilities into degrees of goodness. Analyzing weight in terms of degree of *explanation* isn’t promising, since it wrongly predicts that nonnormative reasons would also have weights.

Sixth, REG offers a natural explanation of the *nonadditivity* of the weights of (some) reasons.<sup>10</sup> As is often observed, in figuring out what we ought to do we cannot always add up the weights of all the reasons we have, because some pairs of reasons are normatively “overlapping”. For these pairs, adding their weights would involve unacceptable double counting. For example, *that rain was forecast* is a reason to take your umbrella, and it is a different fact—and therefore, presumably, a different reason—from the fact *that it is raining*. Each fact might be a sufficient reason, by itself, to take your umbrella, but their weights cannot be added. Suppose you observe that it is raining, then realize you have a marginally stronger reason *not* to take your umbrella (e.g. it will take a few minutes to find, making you late to work). It would be absurd then to resolve your deliberations in favor of taking your umbrella on the basis that additionally, *rain was forecast*. The fact that rain was forecast has no weight in this scenario over and above the weight of the fact that it is raining.<sup>11</sup> REG explains this, because (i) it accounts for the weights of reasons by the degree and kind of goodness they explain, and (ii) we can have non-competing correct explanations of the same thing. The fact that it is raining and the fact that rain was forecast are both explanations of the same kind of value that taking your umbrella has, involving your staying dry. Since REG explains normative weight in terms of degrees of value, it explains why two reasons involving the same value will have weights that are not additive.

Finally, REG *reconciles central platitudes about rational agency* which otherwise seem in tension. On one hand, it is a platitude that a rational agent tries to choose the best available option. On the other hand, it is a platitude that a rational agent chooses the option most supported by her reasons. It is therefore natural to interdefine goodness and reasons, although this consideration is neutral between analyzing reasons in terms of goodness (REG), or goodness in terms of reasons (“buck-passing”).

Observing these seven virtues should be enough to earn REG a hearing, so I turn now to the first of our five puzzles.

## 2. Reasons for Attitudes of the “Right Kind”: The Puzzle

A well-known obstacle to value-based theories of reasons such as REG comes from the need to account for normative reasons for *attitudes*, including both doxastic attitudes like belief (“epistemic reasons”) and emotional attitudes like fear, hope, shame, pride, guilt, and anger. Here a distinction is commonly drawn between reasons of the “right kind” and of the “wrong kind”.<sup>12</sup> Reasons of the wrong kind (“state-given”, “pragmatic” reasons) are those involving beneficial consequences of having the attitude; for example, the fact that somebody will reward or refrain from punishing you if you believe, fear, hope, etc. Reasons of the right kind (“object-given”), by contrast, involve the intrinsic *fittingness* of the attitude to its object, rather than beneficial consequences of having the attitude. The most discussed

---

<sup>10</sup> Maguire 2016, Maguire & Snedegar ms. Brunero’s (2013: 810-11) objection overlooks this feature.

<sup>11</sup> The reverse is not true, because the probability of getting wet without an umbrella conditional on its raining is higher than its probability conditional on rain merely being forecast.

<sup>12</sup> Rabinowicz & Rønnow-Rasmussen 2004. These labels derive from the attempt to analyze value of various kinds in terms of reasons, the issue being whether a reason is of the right kind to ground the relevant evaluative property. For example, the fact that someone will pay you to believe, fear, or be ashamed that *p* isn’t the right kind of thing to make it that *p* is credible, fearsome, or shameful.

example is epistemic reasons for belief that  $p$ , which according to evidentialism consist exclusively in evidence that  $p$  is true, independent of positive consequences of believing that  $p$ .<sup>13</sup> However, the right kind of reasons for emotional attitudes like fear, hope, and shame pattern the same way, depending on the attitude's fit to its object rather than its beneficial consequences.<sup>14</sup>

On face value, REG provides only an account of “wrong kind” reasons (in terms of the attitude's value), and doesn't extend to “right kind” reasons. So if there are any “right kind” reasons—as virtually everybody agrees—then REG appears at best seriously incomplete.<sup>15</sup> Some philosophers claim more boldly that “reasons of the wrong kind” are not genuinely reasons at all.<sup>16</sup> On one line of thought, reasons by definition are things that agents can rationally form attitudes in response to, and we can't rationally believe, fear, hope, etc. on the basis of considerations about the value of having those attitudes. It is suggested that “reasons for action” are strictly speaking reasons for *intention*, a special case of “right kind” reasons for attitudes.<sup>17</sup> At worst, therefore, there may be no normative reasons corresponding to REG at all. This problem motivates rival theories of normative reasons in terms of fittingness,<sup>18</sup> and is a challenge for REG along with many other theories.

## 2.1 Goodness as End-Relational

What REG predicts about what reasons there are depends on how we should understand the relevant concept of goodness. I will show that when this is understood in a certain way, sketched in this section, the “right kind of reasons” puzzle—along with our other four puzzles—is amenable to a straightforward and systematic solution.

While the concept of goodness does the heavy lifting in this paper, we will also need a better than pretheoretic grasp of the similarly contested concept of *explanation*. The relevant sense of ‘explanation’ here is that of an *explanans*, or what plays a “because” role in a true statement of the form ‘ $p$  because  $q$ ’, and not that of an act (type or token) of explaining. I assume a fairly orthodox model theory, as follows. Start with the following rough definition of the technical notion of a “complete explanation”, in the tradition of the Deductive Nomological (DN) theory<sup>19</sup>:  $p$  is a *complete explanation* of  $q$  in case  $p$  is a set of true propositions that logically entails  $q$  but doesn't include  $q$ . Logical entailment may seem too strong here, but  $p$  should be understood as including any relevant conceptual truths or essential definitions, occupying the role played by scientific laws in Hempel's theory of scientific explanation.<sup>20</sup>

---

<sup>13</sup> E.g. Kelly 2003, Shah 2006, Brunero 2013.

<sup>14</sup> E.g. Rabinowicz & Rønnow-Rasmussen 2004, Skorupski 2007, Gardner 2008, Raz 2009: 47-8, 2011: 48, Heuer 2010: 177-8, Way 2012, 2013.

<sup>15</sup> E.g. Gardner 2008, Raz 2011: 42, Way 2013, Brunero 2013; cf. Maguire 2016, who as a proponent of value-based reasons draws the radical conclusion that “right kind” reasons aren't genuinely reasons.

<sup>16</sup> E.g. Skorupski 2007: 10f, Way 2012, Parfit 2011a: 431-2.

<sup>17</sup> E.g. Scanlon 1998: 21.

<sup>18</sup> E.g. Yetter Chappell 2012, McHugh & Way 2016, Howard 2017.

<sup>19</sup> Hempel 1965.

<sup>20</sup> DN theories face well-known problems; see Woodward 2014. However, the main problems will be moot for our purposes. (1) The problem of *statistical explanations* doesn't arise because the proposed normative explanations (explanantia) won't be statistical. They will emerge to be explanations of probabilities, i.e. of statistical explananda, conforming with standard DN solutions. (2) Although this DN theory is admittedly too broad, the problems of *irrelevancies* and *asymmetries* don't arise

However, our concern will be with ordinary statements about “reasons”, which generally address partial rather than complete explanations. This ordinary use is context-sensitive, so that “explanations” and “reasons” are identified relative to a background set of assumptions or information, B.<sup>21</sup> For example, the fact that *I took my umbrella* qualifies as an explanation or reason why I stayed dry in the rain only against a set of background assumptions including that if I am in the rain with an umbrella then I use it, and that using an umbrella in the rain keeps one dry. Here is the model semantics I’ll adopt:

A fact/true proposition *that p* is an explanation (reason) why *q*, relative to B, in case (i) *p* and B together entail *q*, and (ii) neither *p* nor B by itself entails *q*.<sup>22</sup>

By default, the content of B is determined by a function from context. At a rough approximation, it can be identified as the conversational common ground or what the audience is able to take for granted, except that to identify *p* as an explanation of *q* functions in discourse (by accommodation) to exclude both *p* and *q* from B, even if they were already in the common ground. So if a sufficient ground for *q* is already known, we can pick out any element of that ground and correctly identify it as an explanation why *q*. Importantly for our purposes, however, this information-parameter can also be explicitly shifted, as by qualifying an explanation as being “for *S*”, or subject-relative. To say that *p* is an “explanation/reason for *S*” why *q* is, in general,<sup>23</sup> to indicate that *p* is an explanation relative to *S*’s (other) information why *q*. For example, the fact that I took my umbrella is not an explanation why I stayed dry in the rain for somebody who doesn’t know what umbrellas are.<sup>24</sup>

Now consider the concept of goodness. While discussion of our five puzzles, and resistance to REG more generally, has overwhelmingly proceeded as if goodness were a single gradable property that things either have or don’t, this assumption is rejected by most recent work on the semantics of the adjective ‘good’, which finds it also to be a context-sensitive term.<sup>25</sup> There is no property of goodness-simpliciter because, as Judy Thomson puts it, to be “good” is always to be good *in some way*.

In response to a claim that something is “good”, for example, we can ask “good for whom?”, “good for what?”, “good to do what with?” ‘Good’ is an adjective that we often qualify in a wide range of ways and with different kinds of prepositional phrases: *good for me/ you/ diabetics/ people/ dogs/ trees/ cars, good to wear/ have/ see*, etc. Even when we say that something is “good”, sans phrase, our intended meaning can at least often be precisified by identifying such qualifiers as assumed or implicit. For example, saying ‘Chocolate is good’ is naturally understood as claiming that chocolate is good *to eat/taste*, rather than (e.g.) *to wear*, and assertion of ‘Winning the lottery is good’ is naturally

---

because the explanantia I’ll propose will be clearly both relevant and metaphysically prior to their explananda, given my reductive account of goodness.

<sup>21</sup> E.g. Raz 1990, 2011: 17, Gardner & Macklem 2004, Heuer 2006; cf. Swanson 2010 on causal explanations.

<sup>22</sup> The second condition is motivated partly by the intuition that facts don’t explain themselves. It may be too strong, since it doesn’t recognize a complete explanation as an explanation relative to an empty background. It stands in for some more nuanced condition requiring a genuine explanation to be informative somehow (e.g. *p* being metaphysically more fundamental than *q*).

<sup>23</sup> This isn’t always the right interpretation for *normative* “reasons for *S*”, which often require a normative and not purely epistemic relation to *S*; see Finlay 2014: 94f.

<sup>24</sup> However, identifying *p* as an explanation can add elements to B that weren’t in the common ground (by accommodation), as transparently needed for *p* and B to be together sufficient for *q*.

<sup>25</sup> E.g. Ziff 1960, Thomson 1992, 1996, Foot 2001, Szabo 2001, Shanklin 2011, Finlay 2001, 2014, Wolfsdorf 2018.

understood as saying it is good *for the winner*, rather than (e.g.) *for humankind*. While the correct semantics for ‘good’ can’t be settled here,<sup>26</sup> my strategy will be to introduce an analysis motivated by the desideratum of a unifying semantics and syntax underlying the diverse ways the word is used, and then show how this theory enables REG to solve all five puzzles. I’ll build this theory incrementally until we have all the resources we’ll need.

To develop this theory with a concrete example, take the simple sentence,

(1) Umbrellas are good.

As uttered in a particular context, this needs to be understood as elliptical for another, more complex sentence, of a form identified through the following progression:

(1a) Umbrellas are good *to use*.

(1b) Umbrellas are good *for people to use*.

(1c) Umbrellas are good *for staying dry in the rain*, for people to use.

(1d) Umbrellas are good for *people’s staying dry in the rain*, for them to use.

(1e) *It is good for people’s staying dry in the rain*, for them to use *umbrellas*.<sup>27</sup>

In different contexts, (1) could be used as elliptical for indefinitely many other sentences of the syntactic form of (1e),

It is good for  $S_1$ ’s *doing*  $A_1$ , for  $S_2$  to do  $A_2$ .

Consider for example its utterance in response to the question, “What can a painter depict to convey an impression of rain?” Here (1) is naturally interpreted as elliptical for

(1d’) Umbrellas are good *for painters’ conveying an impression of rain*, for them to depict.

(1e’) It is good for *painters’ conveying an impression of rain*, for them to depict umbrellas.

According to this analysis, ‘good’ fundamentally expresses a relation between two propositional arguments, which I’ll call the *object* proposition,  $p$  (e.g. *people use umbrellas*), and the *end* proposition,  $e$  (e.g. *they stay dry in the rain*).

We can now turn from the syntactic question of logical form to the semantic question of meaning: what does it mean to say that  $p$  is good for  $e$ ? A natural answer is: that  $p$ ’s obtaining *raises the probability* of (or “promotes”)<sup>28</sup>  $e$ ’s obtaining. For example, that *people use umbrellas* raises the probability that *they*

---

<sup>26</sup> I attempt this in Finlay 2014, ch. 2.

<sup>27</sup> The transition to (1e), effecting an analysis of the goodness of objects into the goodness of states of affairs, interprets the grammatical structure of sentences (1)-(1d) in terms of *subject-movement*, whereby we satisfy the English grammatical requirement of a subject-term for sentences whose verbs are propositional operators by moving a noun-phrase from the embedded, complement sentence (‘People use umbrellas’) to subject-position. The telltale trace of this movement is the verb left without an object in (1a)-(1d): ‘...to use \_\_\_\_’. For discussion of subject-movement and ‘good’, see Shanklin 2011, Finlay 2014: 29f.

<sup>28</sup> The equivalence of “promotion” and probability-raising is challenged for omitting an essential causal relation (see Behrends & DiPaolo 2011 and follow-up literature; cf. Nagel 1970: 51); this is related to the choice between evidential and causal decision theory. I assume it (and, implicitly, evidential decision theory) as a rough gloss, with probability-raising (as analyzed in n. 45) as the preferred term; see Finlay 2014: 39f.

*stay dry in the rain*, and that *painters depict umbrellas* raises the probability that *they convey an impression of rain*. According to this end-relational theory (ERT), goodness involves a relation to an end or outcome. This sketch leaves many important questions and objections unaddressed, of course.<sup>29</sup> Some of these involve how to understand the relevant concept of probability, a question to which I'll return.

Combining REG with ERT yields the following analysis:

REG+ERT (first pass): To say that R is a "reason" for S to do A is to say, of some end *e*, that R is an explanation why it would be good/probability-raising for *e*, if S does A.

Not just *any* end or outcome licenses a claim that something is a normative reason for action or attitude, of course. On this contextualist view, something is correctly said to be "good" (sans phrase) in a particular context only if it is good relative to an end that is in some way salient or privileged in that context, and so likewise a fact is correctly described as "a reason" (sans phrase) in a context only if it is an explanation of goodness relative to a salient or privileged end.<sup>30</sup> This end-relativity of goodness is what will enable REG to solve the "right kind" reasons puzzle for attitudes, as I'll now explain.<sup>31</sup>

## 2.2 "Right Kind" of Reasons for Attitudes: A Solution

Plausibly, an end is privileged in a context generally in virtue of being a desired or preferred outcome for some salient subjects, who are normally some combination of the speaker, audience, and the agent under discussion.<sup>32</sup> We can thereby explain the intuitions of Humean internalists about reasons as the result of privileging the *agent's* preferences for outcomes (e.g. Hitler may well have had "no reasons" to refrain from genocide, assuming that none of his preferred ends were threatened by it), while explaining the intuitions of externalists, and the categorical nature of moral claims, as the result of privileging the *speaker's* or *audience's* preferences for outcomes (*of course* Hitler had reasons to refrain genocide, given the harm it inflicted on innocent people!)<sup>33</sup>

This sharpens rather than resolves our first puzzle. Something is a reason of the "right kind" for an attitude if and only if it makes the attitude fitting, but raising the probability of an end desired by the speaker, audience, or agent is neither necessary nor sufficient for fittingness. However, when talking about reasons for *attitudes* there is a competing source of salience for ends, in the reference to the attitudes themselves. These are attitudes that are commonly said to have subagential, "constitutive" ends of their own, which need not be ends desired or intended by anybody. It is commonly said, for example, that the constitutive end of belief is *truth* (or knowledge). How to precisify this idea of

---

<sup>29</sup> For more details and defense, see Finlay 2014, ch. 2. As an instrumental analysis one obvious objection is that it doesn't account for final goodness. For a reply see Finlay 2014: ch. 7.

<sup>30</sup> Barry Maguire is therefore incorrect to distinguish our versions of REG on the grounds that my view "places no constraints on ends" (2016: n3). I agree with him that a "reason" to do A (simpliciter) is an explanation why doing A would promote an outcome that is "good" (simpliciter). However, some opponents of REG (e.g. Raz 2011, ch. 11) may observe that REG+ERT, with its relativistic analysis of value, was never their intended target.

<sup>31</sup> See also Finlay 2014: 100-3.

<sup>32</sup> See Harman 1996:15-16, 59f and Finlay 2014: chs. 5-8.

<sup>33</sup> See especially Harman 1996: 59f, Finlay 2006, 2014: ch. 7.

“constitutive ends” is controversial, but for our purposes this doesn’t matter.<sup>34</sup> All we need is that talking about “belief” is sufficient, in normal contexts, to make salient an end like truth.

Consider for example the simple view that our doxastic attitudes or faculties constitutively aim at *thereby*<sup>35</sup> believing that  $p$  if  $p$  is true and not believing that  $p$  if  $p$  is false. On REG+ERT, calling  $R$  a “reason to believe that  $p$ ” would then have a natural interpretation as follows:

*Truth-Relative Reasons:*  $R$  is an explanation (for  $S$ ) why  $S$ ’s believing that  $p$  would be good for/ raise the probability of  $S$ ’s *thereby believing that  $p$  iff  $p$  is true.*

In general, an explanation for a subject of why believing  $p$  raises the probability that (one believes  $p$  iff  $p$ ) will be something that itself raises the probability that  $p$  over *not- $p$* .<sup>36</sup> In other words, it will be *evidence* for  $p$ , on an orthodox definition of evidence. So this interpretation identifies “reasons for belief” generally with evidence, just as evidentialist opponents of REG insist (although not *all* evidence; see note 38.)

Does REG+ERT yield the correct analysis of “right kind” reasons for belief? I cannot answer this question here, as there are at least two confounding factors. First, while *Truth-Relative Reasons* certainly faces objections and potential counterexamples, there is a lack of philosophical consensus on exactly when something is an epistemic reason. Second, REG+ERT is not committed to this precise analysis as it rather provides a template that can be applied to rival (evidentialist and nonevidentialist) conceptions of the constitutive end(s) of belief: for example, the end of *believing  $p$  iff the probability of  $p$  is greater than (some value)  $n$* .<sup>37</sup> So to ultimately judge its success one might need first to resolve the controversy over the constitutive or proper end of belief. But what we have found is that when combined with ERT, REG does support a reading of “reasons for belief” that depends on a subject’s evidence and not at all on her desires or interests, and therefore that it has the resources in principle to accommodate “right kind” reasons for belief.<sup>38</sup>

---

<sup>34</sup> On one view it is constitutive of beliefs that they can only be formed in response to considerations that seem to bear on their truth (=evidence). See especially Williams 1973, Shah & Velleman 2005, Hieronymi 2006.

<sup>35</sup> Including ‘thereby’ avoids a special case of “wrong kind of reasons” problem, where believing (incorrectly) that  $p$  at  $t_1$  raises the probability of believing (correctly) that *not- $p$*  at  $t_2$ . This is not ad hoc, as the constitutive end of a belief token (as of all these attitudes) is local to that token itself.

<sup>36</sup> Proof:  $\text{pr}(Bp \leftrightarrow p)$  is equivalent to  $\text{pr}((Bp \vee \sim p) \wedge (\sim Bp \vee p))$ . When  $\sim Bp$ , the second conjunct’s probability is 1, and so the probability of the conjunction is just the probability that  $(Bp \vee \sim p)$ , i.e.  $\text{pr}(\sim p)$ . When  $Bp$ , the first conjunct’s probability is 1, so the probability of the conjunction is just the probability that  $(\sim Bp \vee p)$ ; i.e.  $\text{pr}(p)$ . Therefore,  $Bp$  raises the probability that  $Bp \leftrightarrow p$  only if  $p$  is more likely than  $\sim p$ . Since an “explanation” of  $x$  is something without which (ceteris paribus)  $x$  would not be true, any explanation why  $Bp$  raises the probability that  $Bp \leftrightarrow p$  will itself be something that makes it the case that  $\text{pr}(p) > \text{pr}(\sim p)$ —except when  $\text{pr}(p)$  is itself dependent on  $\text{pr}(Bp)$ , as with self-fulfilling beliefs.

Is this exception a problem for REG+ERT? Consider (R) *that it is within your power to make  $p$  true or false*; e.g. suppose  $p$  concerns your own future actions, and that believing  $p$  raises the probability that you believe  $p$  iff  $p$  because it raises the probability that you act so that  $p$ .  $R$  is plausibly not evidence for  $p$  prior to your forming the belief that  $p$ ; might it nonetheless be a reason to believe  $p$ , as *Truth-Relative Reasons* appears to predict? The counterintuitiveness of this might be softened by observing that if so it is equally a reason to believe  $\sim p$ , and hence to not believe  $p$ . Any such reasons would cancel each other out, making it unhelpful to mention them. Alternatively, such cases could be excluded by replacing  $S$ ’s believing  $p$  in the analysis with a *generic subject’s* believing  $p$  with  $S$ ’s information, making  $p$  probabilistically independent of  $Bp$ .

<sup>37</sup> Thanks to Jaime Castillo Gamboa for discussion.

<sup>38</sup> One problem arises from the template itself rather than its application to any particular end. REG+ERT may seem to deny the possibility of *outweighed* reasons for belief (see also note 7), because plausibly there is an explanation why believing  $p$  raises

The same template can be applied systematically to generate accounts of right-kind reasons for other attitudes. It is commonly said, for example, that desire aims at *the good*. So REG+ERT predicts that talk about a “reason to desire” will have a natural interpretation as *an explanation why desiring that p would be good for thereby desiring that p iff p is good*. Other kinds of (“state-given”) benefits, like monetary rewards for desiring, will not provide reasons to desire in this sense.<sup>39</sup> Similarly, intention plausibly aims, approximately, at what is *all-things-considered best to do*, in which case talk about a “reason to intend” will have a natural interpretation as an explanation why it would be good, for thereby intending to do A iff A is all-things-considered best to do, to intend to do A. This analysis blocks familiar “toxin puzzle” cases: the fact that intending to drink poison will yield a monetary reward is not an explanation why intending to drink poison would be good *for thereby intending to drink poison iff drinking poison is all-things-considered best to do*, and therefore is not a “right-kind” reason for this intention. Similarly, fear aims approximately at the dangerous, hope aims at that which is both possible and good, shame aims at what is to one’s discredit, and amusement aims at the funny. So we can understand a “reason to fear (hope, be ashamed, be amused) that p” as an explanation why it would be good, for thereby fearing (hoping, being ashamed, being amused) that p iff p is dangerous (both possible and good, to one’s discredit, funny) to fear (hope, be ashamed, be amused) that p.

REG+ERT systematically generates results in all these cases that preserve the features of “right kind” reasons with which it is alleged that REG is incompatible. An added bonus is that it also offers an explanatory analysis of fittingness, which some philosophers claim to be an unanalyzable primitive:<sup>40</sup> for an attitude to “fit” its object is for that attitude to realize or promote its constitutive end when directed at that object. Reasons of the “right kind” for attitudes are facts that explain why an attitude fits its object in this way. Finally, REG+ERT also explains the conflict in intuitions about reasons for attitudes. If we privilege the salience of attitudes’ constitutive ends over other ends such as those desired by the subjects or speakers, then it is correct to deny that “wrong-kind” reasons are appropriately cited as reasons at all. We might do so, perhaps, on the ground that agents cannot (at least rationally) form attitudes in response to any but “right-kind” reasons, making it conversationally pointless or misleading to mention any other kind. But if we rather privilege (e.g.) subjects’ desired ends over those constitutive ends, then in those contexts we rightly recognize pragmatic reasons as such, and perhaps even deny that “right-kind” reasons are reasons at all.

---

the end’s probability only if believing *p* does raise that probability. To illustrate for *Truth-Relative Reasons*, evidence raising the probability of *p* from 0.1 to 0.2 doesn’t meet this condition, because believing *p* when one’s evidence puts its probability under 0.5 makes it *less* likely that you believe *p* iff *p*. Some epistemologists indeed deny that such evidence is a reason to believe (e.g. Roush 2005: 158), but some readers may find it too radical an implication. My preferred solution is to accommodate such reasons in terms of *ceteris paribus* goodness, or probability-raising relative to a restricted information-base excluding one’s other evidence. (The information-relativity of goodness is explored in section 3.) Work is needed to show this can be done nonarbitrarily, but partners in guilt are plentiful since standard probabilistic analyses of evidence need similar moves to account for redundant evidence, received when the probability of *p* is already 1.

<sup>39</sup> It might be objected that this analysis requires a non-instrumentalist theory of goodness incompatible with ERT. However, *instrumental* desire is plausibly aimed at *instrumental* goodness as analyzed by ERT, and we can consistently deny both (i) that *intrinsic* desire aims at the good in any nontrivial sense (in which goodness is prior to desire), and (ii) following Hume, that there are any “right-kind” reasons for intrinsic desires. See Finlay 2014: 197f for an ERT-based analysis of final goodness.

<sup>40</sup> E.g. Yetter Chappell 2012, Howard 2017.

### 3. From End-Relativity to Information-Relativity

While the hypothesis that goodness is end-relative directly provided a solution to our first puzzle, the solutions to our other four puzzles are less direct, and require exploring a further way in which ERT implies that the concept of goodness is context-sensitive. Above I analyzed goodness-for-an-end in terms of raising the probability of the end. This prompts further questions, because there are different senses of probability; for example, we can distinguish between objective and subjective probabilities.

While the analysis of the concept of probability is itself difficult and controversial, a powerful, flexible, and unifying approach is to interpret it also as context-sensitive, and involving a relation to an information-base which can vary from context to context. On this view, claims about probabilities are made and evaluated against an information-base assumed as a conversational background, as in orthodox linguistic treatments of natural language modals.<sup>41</sup> By default this background can be identified as the common ground in the conversation. This makes claims about probability (and goodness) relative to the same kind of information parameter as claims about explanation (as analyzed in section 2.1). For simplicity, I will assume by default that in statements about normative reasons, explanation and probability/ goodness are relativized to the same background information.<sup>42</sup>

This further argument of ‘good’ can be made explicit or shifted with a ‘given...’ or ‘relative to...’ clause; e.g. ‘Relative to *what we know*, the probability of staying dry is higher if you take an umbrella’; ‘Given *what the weather forecast says*, it would be good to take an umbrella’. It will prove fruitful to interpret this argument intensionally—again following the orthodox treatment of modals, identifiable with a particular description or property *b* (e.g. *what we know*). In a context *w*, this intension *b* determines as its extension a set of propositions *b(w)* (informally, a partial sketch of the world), relative to which the relevant probabilities are determined. The rationale for this complication is that it allows us to distinguish talk about what is probable, good, or a reason *relative to what S<sub>1</sub> believes* from talk about what is probable, good, or a reason *relative to what S<sub>2</sub> believes*, even if S<sub>1</sub> and S<sub>2</sub> have exactly the same beliefs. This will be important below.

We now reach the final form of the proposed analysis of normative reasons statements:

REG+ERT: To say in a world *w* that R is a “reason” for S to do A is to say, of some end *e* and information-base *b*, that R is an explanation in *w* why *given b* it would be good/probability-raising for *e*, if S does A.<sup>43</sup>

We can then analyze objective probabilities as relative to an information-base defined in some objective way (e.g. *the world’s state at time t*), and subjective probabilities as relative to an information-base defined in terms of information available to some subject.<sup>44</sup> The significance for our remaining puzzles is

---

<sup>41</sup> E.g. Kratzer 2012.

<sup>42</sup> With one important exception: the information-base for the goodness must include the “reason” R itself (as helping to determine the goodness), which is excluded from the explanation’s information-base.

<sup>43</sup> Our analysis suggests that ‘a reason for S to do A’ is ambiguous between ‘a reason [for S<sub>1</sub>], for S<sub>2</sub> to do A’, and ‘a reason for S<sub>1</sub>, [for S<sub>2</sub>] to do A’; see Finlay 2014: 94f. Here I avoid complications by assuming the identity of S<sub>1</sub> and S<sub>2</sub>.

<sup>44</sup> For discussion see Finlay 2014: 39f. Note that “subjective probabilities” here are not psychological *credences*, but measures of a possibility-space defined by a subject’s beliefs or evidence, which can diverge from her credences.

that this generates different information-relative senses of goodness. For example, we can distinguish between objective goodness, as raising the objective probability of an end, and subjective goodness, as raising the subjective probability of the end for some subject.<sup>45</sup> We can further discriminate between statements about subjective goodness relative to the agent's information, and statements about subjective goodness relative to the speaker's or judge's information. The rest of the paper demonstrates how this information-relativity of goodness provides solutions to the four remaining puzzles.

#### 4. Normative Facts as Reasons: The Puzzle

A second puzzle for REG involves cases of normative facts that are themselves normative reasons. For example, the fact that *doing A would be good* is arguably itself a reason to do A. If so, this looks like a problem for REG, which would then imply that the fact that *doing A would be good* is itself an explanation why doing A would be good.<sup>46</sup> But plausibly nothing can be literally self-explanatory; i.e. no proposition  $p$  is such that  $p$  is an explanation why  $p$ . So it appears there are some reasons to do A that aren't explanations why it would be good to do A.

The following scenario provides such a case:

*Posting Money.* Ingrid's trusted, reliable friend Hannah tells her that it would be good for her to put \$500 in an envelope and post it to a particular, unfamiliar address N in another city. She doesn't explain why this would be good.

Ingrid intuitively now has a reason to post money to N, whereas previously she was unaware of any reason to do so. At least, it would be reasonable for her to post money, and plausibly an action is reasonable only if the agent has a reason for it. (The relationship between "there being" a reason for an agent, and that agent "having" a reason, is debated. Here I assume, fairly uncontroversially, that the possessive constructions generally indicate the agent's awareness.)<sup>47</sup> Of course, one reason Ingrid arguably now has is the fact that

HS: Hannah told her that posting money to N would be good.

(We'll explore a problem involving this kind of reason in section 5). But the normative status of HS as a reason to post money surely depends on whether what Hannah tells her, that

PG: Posting money to N would be good,

---

<sup>45</sup> I shall assume that  $p$  "raises" the probability of  $e$ , relative to  $b$ , iff the probability of  $e$  given  $b$  and  $p$  is higher than the probability of  $e$  given  $b$  and  $\sim p$  (Finlay 2006, 2014: 39f).

<sup>46</sup> Versions of this problem have been raised by Daan Evers (2010: 412) and Mike Ridge (correspondence); my treatment in this paper expands on Finlay 2014: 94n.

<sup>47</sup> E.g. Schroeder 2008; Williams 1979 takes them rather as (defeasibly) suggestive of an agent's motivationally "internal" reasons, while Gibbons 2010 denies that the difference is significant. I distinguish five different readings in Finlay 2014: 103-114.

would itself be a reason to post money if true. Ingrid wouldn't recognize HS as a reason to post money if she were convinced that Hannah's claim (PG) is false, for example, and *being told that p* is generally only a reason to do A if *p* would itself be a reason to do A, if true.<sup>48</sup>

It is controversial whether normative facts like PG are ever normative reasons to act. This is denied especially by "buck-passing" theorists who define the goodness of an action or attitude by the existence of reasons for it.<sup>49</sup> If the buck-passers are right then there is no genuine problem for REG here, but I believe they are mistaken. The intuitive case for denying that facts like PG can be reasons comes from the observation that offering normative facts as reasons is often unhelpful and inappropriate. Consider:

*Posting Money—Explained:* Asked why it would be good for Ingrid to post money to N, Hannah explains that N is the address of Ingrid's long-lost brother Francis, with whom Ingrid would love to be reunited. Hannah further explains that Francis is contemplating relocating to the city where Ingrid happens to live, a move that depends on an otherwise unlikely influx of cash.

Ingrid now has a reason for posting money to N that can be described (in short) as the fact that

FA: N is her brother Francis's address.

It would be strange, now that Ingrid knows FA, to tell her that she has an additional reason to post money to N, which is the fact PG that *doing so would be good*.

As others have pointed out, however, this can plausibly be explained by the nonadditivity of the weights of some reasons.<sup>50</sup> Since PG, the goodness of posting money to N, is itself explained by FA, the fact that N is Francis's address, PG has no normative weight over that of FA. This goes some way toward explaining why presenting PG to Ingrid as an "additional reason" is inappropriate, although I'll suggest below that REG+ERT can improve on it.

That the normative fact PG can still have the normative weight of a reason is supported by the following consideration. Suppose that in the original *Posting Money* scenario (where Ingrid doesn't know FA) she has complete trust in Hannah and therefore comes to believe confidently that posting money to N would be good, without knowing why. If asked, "What reason do you have to post money?" she could appropriately reply, "My reason is that it is good."<sup>51</sup> All else equal, an agent who believes that doing A would be good, without any further information, can rationally decide to do A on this basis, so it also seems such facts can play the characteristic psychological role of normative reasons in our

---

<sup>48</sup> An exception: when being told that *p* functions as a sign of something else that is a reason to do A. But *Posting Money* isn't a case of this kind.

<sup>49</sup> E.g. Scanlon 1998, Stratton-Lake & Hooker 2006: 161; see also Dancy 2004: 16. For the case against, see Crisp 2005, Väyrynen 2006, Schroeder 2009a, Heuer 2006, 2010b.

<sup>50</sup> Schroeder 2009a, Heuer 2010b.

<sup>51</sup> Citing HS might be more natural. However, (i) this may pragmatically suggest a smidgeon of doubt about Hannah's reliability, and (ii) we can suppose that Ingrid forgets HS/ how she learned of PG.

deliberations.<sup>52</sup> So it remains a problem for REG if PG cannot be an explanation why acting would be good.

Similar considerations apply to normative facts concerning the *existence of reasons*. Suppose Hannah tells Ingrid simply that

RE: There is a reason for Ingrid to post money to N,

and so Ingrid comes to believe this. It seems Ingrid now has a reason to post money, and that RE is one of the reasons she now has. At least, she could rationally decide to post money on this basis. This also looks embarrassing for REG, which analyzes RE as saying that there is an explanation why doing A would be good. The bare existential fact that *there is an explanation why doing A would be good* doesn't seem as if it could itself be an explanation why doing A would be good. These problems for REG and other explanation theories encourage rival theories of normative reasons, such as the view that they are the premises of sound practical reasoning.<sup>53</sup>

#### 4.1 Normative Facts as Reasons: A Solution

A solution to this puzzle emerges from REG+ERT's feature of information-relativity. We saw in section 3 that the relativity of 'good' to an informational background generates different senses of goodness. While perhaps nothing can explain itself, I'll now show that the goodness<sub>1</sub> of S doing A can be an explanation of the goodness<sub>2</sub> of S doing A—and indeed is, in precisely the right circumstances—through a context-shift in the information-base *b*.<sup>54</sup>

Consider our original scenario, *Posting Money*. ERT tells us that the normative information that Hannah provides Ingrid, (PG) *that it would be good for her to post money to N*, must be interpreted as involving at least two implicit arguments for 'good': an end *e* and an information-base *b*. While there are significant issues surrounding the choice of *e*, I will assume that *e* is picked out under the description *an outcome Hannah has in mind that Ingrid would prefer if fully informed*.<sup>55</sup> The salient information-base is presumably *Hannah's information*, as the context is transparently one where Hannah is advising Ingrid from her position of fuller information. So we should interpret PG along the following lines:

PG-H: The probability relative to *Hannah's information* (*b<sub>H</sub>*), of *an outcome Hannah has in mind that Ingrid would prefer if fully informed* (*e*), is higher if Ingrid posts money to N (*p*) than if not ( $\sim p$ ).

Formally,

---

<sup>52</sup> See Heuer 2010b: 149. Might FA still be the relevant reason for explaining Ingrid's rationality? Although she doesn't know FA she still knows that there is some fact that is a reason to post money. But this overlooks that Hannah has *given* Ingrid a reason to post money, which she now has (although as discussed below the fact *that there is a reason to do A* is itself plausibly a reason to do A).

<sup>53</sup> E.g. Williams 1979, Setiya 2007, Raz 2011: 23. Wedgwood 2015 claims we should recognize normative reasons of both kinds, rejecting the possibility of a unified account.

<sup>54</sup> Some cases might also be resolved by appeal to end-relativity. For example, that doing A would be good *for my children's welfare* is an explanation why doing A would be good *for ends I care about*; cf. Evers 2010: 412.

<sup>55</sup> See Finlay 2014: ch. 5, 6.

PG-H:  $\text{pr}(e | b_h \wedge p) > \text{pr}(e | b_h \wedge \sim p)$ .

What we need to know is whether PG-H is correctly identified as a reason for Ingrid to post money. According to REG+ERT, this is a question of whether PG-H is an explanation for Ingrid why her posting money would raise the probability of  $e$  relative to the salient information-base. What information-base is this? Since we're talking about "reasons for Ingrid", the most natural candidate is  $b$ : *Ingrid's information* (including supplementation with what Hannah tells her); i.e. information Ingrid could take into account in her deliberations. The answer to our question, I'll argue, is positive. The fact that posting money would raise the probability of  $e$  relative to *Hannah's information* is indeed an explanation for Ingrid why posting money would raise the probability of  $e$  relative to *Ingrid's information*.

While I believe that this claim is intuitively compelling, arguing for it will require some work, which I'll tackle in two steps. Readers satisfied with the intuitive claim can skip to the section's end. The first step is to establish the truth of the explanandum, that posting money is indeed good relative to Ingrid's information. The second is to establish that PG-H indeed stands in the explanans relation to that explanandum.

*Step 1: Posting money would be good relative to Ingrid's information.*

Unfortunately, a full demonstration of the explanandum turns out to be extremely difficult.<sup>56</sup> But we can construct a strong presumptive case, as follows. Ingrid's information consists of (i) her background information, relative to which posting money isn't good for  $e$ ; (ii) the information that Hannah provides her, i.e. that there is a fuller body of information, relative to which posting money is good for  $e$ ; and importantly (iii) that she has no other relevant information. Now consider the intuitive and widely-accepted Principle of Reflection:<sup>57</sup>

*Reflection:* If (i) a subject knows at time  $t_1$  that at a later time  $t_2$  with strictly more information and no cognitive disfunction her credence in a proposition  $p$  will be  $n$ , and (ii) she does not know that her credence in  $p$  at some other later time  $t_3$  will be something other than  $n$  on the basis of even fuller information,<sup>58</sup> then her credence in  $p$  at  $t_1$  ought also to be  $n$ .

There are three potential obstacles to applying this principle to *Posting Money*, but they seem easily overcome. First, Reflection remains controversial despite its appeal, and has some alleged counterexamples.<sup>59</sup> However, our case doesn't seem to bear any relevant similarity to these. Second, Reflection is formulated in subjective Bayesian terms as a constraint on rational credences, whereas we are working with a different, nonpsychological interpretation of "subjective probabilities". However, we can reasonably assume that the credence a subject ought rationally to have in  $p$  corresponds to the probability of  $p$  relative to her information, and thus interpret Reflection in a non-psychological way.

---

<sup>56</sup> In effect, this would be to prove a version of the Principle of Reflection.

<sup>57</sup> Based on van Fraassen 1984.

<sup>58</sup> Jake Ross points out that conjunct (ii) is both idiosyncratic and redundant, since conjunct (i) and the Principle of Reflection together entail it. I include it nonetheless because its counterpart in the modalized version proposed below is nontrivial.

<sup>59</sup> E.g. the Sleeping Beauty problem (Elga 2000).

Third, *Posting Money* isn't necessarily a scenario in which the subject will herself have the additional information in the future. However, since Ingrid correctly trusts Hannah as an epistemic authority, she correctly believes that *were* she to have Hannah's additional information without cognitive disfunction she would assess the probabilities just as Hannah does.<sup>60</sup> It seems arbitrary to accept the temporal Principle of Reflection but not a modal version, because the difference seems immaterial between (i) what a better-informed Ingrid *will* believe in the future, and (ii) what a better-informed Ingrid *would* believe. We can derive the following version of the principle:

*Modal Reflection:* If (i) a subject knows there is some specific additional information  $I_1$  beyond what she already knows such that the subjective probability of  $p$  relative to the conjunction of  $I_1$  and her (other) existing information is  $n$ , and (ii) she doesn't know of the existence of any further specific information  $I_2$  beyond what she already knows such that the subjective probability of  $p$  relative to the conjunction of  $I_1$ ,  $I_2$ , and her existing information is something other than  $n$ , then the subjective probability of  $p$  relative to her existing information is also  $n$ .<sup>61</sup>

We can conclude that Ingrid's subjective probabilities will be the same as the probabilities that Hannah communicates. Of course, Hannah doesn't communicate precise probabilities relative to her information, only that the probability of  $e$  relative to her information is greater if Ingrid posts money than otherwise. So Ingrid's subjective probability for  $e$  given  $p$  will be a *range* we can roughly identify as the interval with the lower bound (noninclusive) of the probability of  $e$  given the conjunction of Ingrid's information and that she doesn't post money, and the upper bound (inclusive) of 1:<sup>62</sup>

$$\text{pr}(e|\sim p \wedge b_i) < \text{pr}(e|p \wedge b_i) \leq 1$$

*Step 2: PG-H is an explanation why posting money would be good relative to Ingrid's information.*

Recall from section 2.1 our working model of what it means for a fact  $p_1$  to be correctly identified, in a context, as an "explanation" of another fact  $p_2$ : that there is a set of information  $b$  assumed in the context, such that the combination of  $b$  and  $p_1$ —but not  $b$  by itself—entails  $p_2$ . What we therefore need to find is that PG-H (posting money raises the probability of  $e$  relative to Hannah's information) is sufficient to make it the case that posting money raises the probability of  $e$  relative to Ingrid's information, when combined with some further propositions in the relevant background. Since we're

---

<sup>60</sup> I assume here that trusting Hannah as an epistemic authority entails assuming that without cognitive disfunction one would share Hannah's priors.

<sup>61</sup> One often knows that there is *some* (nonspecific) further information or other that, if one learned it, would raise the probability of  $p$ , and similarly *some* further information that would lower the probability of  $p$ . *Modal Reflection* assumes that such information is irrelevant to our subjective probabilities (i.e. balances out) until we learn something specific, such as that *Hannah has additional information on which  $p$  is more likely*. Unless this is correct, I suggest, the original Principle of Reflection couldn't be true either. However, here I am neglecting (due to publication deadlines) expert misgivings from Daniel Greco, who advises that a more robust proof could be built on a principle of expert deference, as offered in Heddon 2015.

<sup>62</sup> Note that the lower bound has to be read *de dicto*/ intensionally, and is extensionally extremely vague. In reality, Ingrid has pragmatic information that further precisifies her subjective probabilities. For Hannah's statement to be helpful, Hannah must believe that the degree to which  $e$  is more likely if Ingrid posts money is sufficiently significant (as weighted by the desirability of  $e$ ) to outweigh the disvalue to Ingrid of parting with her money. Put simply, posting money must be *significantly* good, relative to Hannah's information, to be worth mentioning.

concerned with reasons or explanations *for Ingrid*, we can identify this background *b* as *Ingrid's other relevant information*.

As it happens, we already implicitly accomplished this in the first step above, arguing for the truth of the explanandum from a set of propositions that included PG-H. However, here is the deductive argument laid out explicitly. (In order to simplify what would otherwise be an unwieldy presentation, I assume ERT in using 'good' as shorthand for 'raises the probability of *e*', and 'value' for 'effect on the probability of *e*'). We can identify a minimal set B of specific propositions and general laws from which the explanandum can be deduced from PG-H (Posting money is good relative to Hannah's information), as follows:

- P1. Ingrid's information  $B_i$  includes information about the value of posting money relative to Hannah's information  $B_h$ .
- P2.  $B_i$  includes the information that  $B_h$  is strictly richer than  $B_i$ .
- P3. There is no body of information  $B_x$  such that  $B_x$  is strictly richer than  $B_i$  and  $B_h$ , and  $B_i$  includes information about the value of posting money relative to  $B_x$ .
- L1. Principle of Modal Reflection.

From these premises we can deduce:

- C1.  $B_i$  entails that posting money is good relative to  $B_h$ . (From PG-H, P1).
- C2.  $B_i$  entails that posting money is good relative to a body of information strictly richer than  $B_i$ . (From C1, P2).
- C3. Posting money is good relative to  $B_i$ . (From C2, P3, L1).

We have therefore identified a background relative to which PG-H, the goodness-relative-to-Hannah's-information of posting money, is an explanation of the goodness-relative-to-Ingrid's-information of posting money.<sup>63</sup>

The goodness<sub>1</sub> of an action relative to one body of information  $B_1$  can indeed be an explanation of the goodness<sub>2</sub> of that action relative to another body of information  $B_2$ , and hence (according to REG+ERT) a reason for that action relative to the right context. It is also easily seen that our demonstration involves the right context. In asking whether PG is a reason *for Ingrid* to post money, the most natural interpretation is of an explanation *for Ingrid*, i.e. relative to Ingrid's other information, and thereby something she could recognize as an explanation and take into consideration in her deliberation. Premises P1-P3 are, indeed, precisely the salient items of Ingrid's background information, and *Modal Reflection* (or a principle in the vicinity) is so intuitive that we expect rational agents to conform their assessments of subjective probability to it. Given default assumptions about context, therefore, REG+ERT predicts it is correct to say that PG is a reason for Ingrid to post money.

Importantly, this result was contingent on  $B_1$  being strictly richer than  $B_2$ , as in *Posting Money* and other cases of normative testimony from better-informed advisors, and so it doesn't overgeneralize to the

---

<sup>63</sup> I suggest that any of the facts in B could be appropriately cited as a "reason" for Ingrid to post money, with PG-H taking its place in the background. Readers can test this against their own intuitions.

wrong contexts. REG+ERT correctly doesn't predict that facts about goodness provide agents with reasons (at least, with normative weight over and above their other reasons) when the goodness is relative to equal or poorer information. So in *Posting Money—Explained*, where Ingrid has all Hannah's information including (FA) that N is Francis's address, the fact PG that posting money would be good relative to Hannah's information provides her with no (nonoverlapping) reason for posting money, supervening as it does on her other information. We thereby accommodate and explain the buck-passers' observation that citing PG as a "reason" in such a context would be inappropriate. Relative to a context that includes the background information that explains why posting money is good relative to Hannah's information, the supervening fact PG does not meet our criteria for an explanation of the goodness of posting money, because the background is already sufficient for the explanandum by itself. Rather than explaining denials that something is a reason on the (potentially unsatisfying) grounds that it is a reason with overlapping weight, we can agree that relative to such a background, these facts are indeed not reasons at all.

Enlightened with this understanding of the context-sensitivity of reasons-talk, we can see how buck-passers' arguments against normative facts as reasons proceed by specifying a context where this subvening information is saliently in the common ground. Meanwhile the arguments of their opponents proceed by specifying contexts where the subvening information is excluded. So REG+ERT also has the virtue of diagnosing and explaining the entrenched disagreement over whether normative facts can themselves be reasons: it depends on the relationship between the information-base of those normative facts and the other facts in the salient background (e.g. known by the agent).

This solution extends to the case of facts about reasons themselves being reasons. If Hannah tells Ingrid merely that (RE) *there is a reason for her to post money to N*, we can analyze this as the claim that there is an (undisclosed) explanation (FA) why posting money would raise the probability of an (undisclosed) desired end *e* relative to Hannah's richer information  $b_h$ . This fact RE is itself information which explains why posting money would raise the probability of *e* relative to Ingrid's poorer information  $b_i$ , and therefore according to REG+ERT is appropriately identified as a reason for Ingrid to post money. Because this is again contingent on RE's status as a reason being relative to a poorer informational background than FA's status as a reason, there is no danger of double-counting or an infinite recursion of higher-order reasons with independent normative weights.

## 5. Evidence as Reasons: The Puzzle

A third, well-known puzzle for REG arises from cases where *evidence* intuitively counts as a reason—not for *belief*, as with epistemic reasons, but for *action*. We already observed such a case in *Posting Money*, in which we noted that the fact HS, *that Hannah told Ingrid it would be good to post money to N*, is plausibly a reason Ingrid has to post money. The problem for REG is that this fact HS is apparently not a plausible candidate for an *explanation why* posting money would be good.<sup>64</sup> In general, the fact that

---

<sup>64</sup> Kearns and Star deny this on the grounds that the normative status is overdetermined (2011: 507-8). But an advisor can't *make it the case* that *p* if it is already the case that *p*, just as a flying rock can't make it the case that/be an explanation why a window is broken if the window is already broken before the rock passes through it. They alternatively suggest an action's

somebody tells you that  $p$  is not an explanation why  $p$ . If what an advisor tells you is true, then at least in ordinary cases its truth is prior to her telling you about it. If Ingrid were to ask Hannah, “*Why* would it be good for me to post money to N?”, she would rightly not accept “Because I told you so” as a helpful or informative reply. Ingrid would be asking after the facts that prompted Hannah to make her claim in the first place.

It seems, therefore, that in cases of normative testimony we can have reasons to act that aren’t explanations why acting would be good, contrary to REG. These reasons apparently take the form rather of evidence it would be good for  $S$  to do  $A$ . (Often, evidence that  $p$  is an explanation why  $p$ , since explanations of something are one kind of evidence for it. However, testimonial evidence is a prominent kind of exception, which will be my focus here.) On this basis, some philosophers argue for a general theory of normative reasons as evidence—rather than explanations—of normative status.<sup>65</sup> It is controversial whether this approach can account for all normative reasons,<sup>66</sup> but it is a significant challenge to REG (and other theories) as a unified account if even some reasons take the form of evidence, and not explanations, of goodness.

### 5.1 Evidence as Reasons: A Solution

A solution to this puzzle also emerges from REG+ERT’s distinction between different senses of “goodness” generated by information-relativity. While it may be true that mere testimonial evidence about something cannot be an explanation of that very same thing, testimonial evidence for goodness<sub>1</sub> (relative to background  $b_1$ ) can be an explanation of goodness<sub>2</sub> (relative to background  $b_2$ ) even when it isn’t an explanation of goodness<sub>1</sub>.

The correct analysis of the concept of evidence is controversial, but the following account is plausible at a first pass: a fact  $R$  is *evidence* that  $p$  for a subject  $S$  in case  $R$  is information that raises  $S$ ’s *subjective probability* that  $p$ ; i.e. if  $p$  is more likely given the conjunction of  $R$  and  $S$ ’s other/ background information  $b$  than it would be given  $b$  by itself.<sup>67</sup> The fact that Ingrid’s trusted friend Hannah tells her that posting money would be good is then evidence that posting money would be good (for some undisclosed desired end  $e$ , and relative to Hannah’s information), while obviously not constituting any explanation why it would be good in that very respect. However, we just saw in section 4 that the fact (PG-H) *that posting money would be good relative to Hannah’s information  $b_h$*  is an explanation why posting money would be good *relative to Ingrid’s lesser information  $b_i$* . This can be extended from the fact PG-H itself to the testimonial evidence for that fact, HS. HS is an explanation why Ingrid’s posting

---

normative status always depends on the agent’s information. This would imply, implausibly, that normative testimony (i) isn’t true before/unless offered (contrary to the overdetermination claim), and (ii) makes itself true.

<sup>65</sup> Especially Kearns & Star 2008, 2009, Thomson 2008. In Finlay 2001: 104f I slid between describing reasons as “explanations” and “indications” of normative facts. Kearns and Star’s analysis is in terms of evidence of ought-facts rather than goodness, which I’ll treat as a restricted evidence-of-goodness view (see n. 7). Switching to an unrestricted evidence-of-goodness theory would provide solutions to some objections to their view, including the Buridan’s Ass objection (cf. Kearns & Star 2009: 238), and reasons that weigh both for and against an action (cf. Kearns & Star 2009: 237).

<sup>66</sup> Wedgwood (2015) argues for a disjunctive theory of normative reasons, distinguishing between reasons-as-evidence and reasons-as-explanations.

<sup>67</sup> Given my working definition of *explanation* (as something that nonredundantly entails in conjunction with the background), it immediately follows that every explanation why  $p$  is evidence that  $p$  (relative to an appropriate background). So I agree that every normative reason to do  $A$  is evidence that doing  $A$  would be good (in some way/to some degree).

money would increase the probability of  $e$  relative to Ingrid's lesser information  $b_i$ . (“Why would posting money increase the probability, relative to Ingrid's present information, that  $e$ ?”; “Because Hannah told her it would be good/increase the probability of  $e$ , relative to her fuller information, if she posted money.”) On REG+ERT this just is to say that HS is an explanation why posting money would be good, relative to Ingrid's information.<sup>68</sup>

So even testimonial evidence that it would be good<sub>1</sub> for S to do A can be an explanation why it would be good<sub>2</sub> for S to do A. We merely need an informational context-shift between the “goodness” that is evidenced, and the “goodness” that is explained.<sup>69</sup> This context-shift isn't gratuitous or ad hoc, because it is exactly what we should expect. Trustworthy advisors offer testimony about what is “good” from the standpoint of fuller information than is possessed by their advisees, which is why we naturally understand Hannah's claim as relative to her fuller information  $b_h$ . But agents have to make their decisions based on what is “best” from their own standpoint, which is why we naturally understand talk of “Ingrid's reasons” as relative to her poorer information  $b_i$ . REG+ERT therefore accommodates the fact that evidence for the goodness of an action can be a reason for that action, and also makes intuitively correct predictions about when it is a reason. The phenomenon of evidence-as-reasons, then, actually supports REG+ERT.<sup>70</sup>

## 6. Subjective Reasons

Our final two puzzles involve so-called “subjective” and “motivating” reasons. These puzzles arise from cases where it seems natural to say that an agent has a reason to do A although, due to the agent being mistaken or uninformed somehow, there are apparently no good candidates for a reason—especially construed according to REG as an explanation of goodness. They stem from the apparent double-factivity of explanation claims:  $p$  can only be an explanation why  $q$  if  $p$  and  $q$  are both the case.<sup>71</sup> (This

---

<sup>68</sup> If Ingrid asks Hannah why posting money would be good, she will likely not view “Because I told you so” as a helpful reply, because she most naturally wants to know why it is good relative to Hannah's information, not relative to her own (which is no mystery).

<sup>69</sup> There is a problem for REG, however, in extending this solution beyond idealized cases like *Posting Money*, in which Hannah's testimony is *sufficient* evidence that posting money is good relative to  $b_h$ . Suppose rather that Ingrid knows Hannah to be somewhat unreliable, or to be contradicted by another, more reliable advisor. In such a case of *insufficient* evidence, posting money might in balance *not* be good relative to all Ingrid's information, although HS is still positive evidence for its goodness. Assuming the factivity of explanation, there then could not exist an explanation why posting money is good relative to all Ingrid's information. But HS is still plausibly a reason (“some reason”) for Ingrid to post money. Various maneuvers might be tried here (thanks to Michael Deigan for suggesting *explanation of why doing A is not worse than it is*), but the most promising in my view is to appeal again to goodness *ceteris paribus*, relative to a restricted information-base that excludes the opposing evidence (see note 7). To counter objections that this move is ad hoc, we can observe (i) the norm of charity, whereby audiences accommodate speakers as much as is feasible by allowing shifts in contextual parameters required to preserve truth, and (ii) that such moves seem obligatory in any case for standard analyses of evidence, to accommodate cases of redundant evidence where the subjective probability of  $p$  cannot be raised, being already 1 (or zero).

<sup>70</sup> This solution expands on Finlay 2014: 93-4. I now think that a fact is evidence of goodness relative to richer information iff it is an explanation of goodness relative to some body of poorer information. How then can we adjudicate between reasons-as-explanations and reasons-as-evidence theories if they are extensionally equivalent? The former are favored as accounts of the *concept* of a normative reason by at least the second (linguistic evidence) and fifth (explaining normative weight) virtues catalogued in section 1.

<sup>71</sup> E.g. Raz 2011: 16. A similar problem arises when the explanatory connection is broken (e.g. Gettier cases in epistemology). I believe the solution below for apparent failures of explanandum-factivity extends to these cases.

claim is controversial; we'll discuss it further below). In this section I focus on a problem from explanandum-factivity, while in section 7 I turn to a problem arising from explanans-factivity.

Consider the following variation:

*Posting Money—In Vain*: All the information Ingrid acquires (in *Posting Money—Explained*) is correct, but there's a further, unknown fact: her brother is about to meet somebody and fall in love, which will end all his thoughts about moving to the city where Ingrid lives. If he receives money in the post, he will just blow it on a boozy party, which he will later regret.

In this case there is apparently nothing actually good about posting money. Given explanandum-factivity, this would imply that there can be no genuine explanations in this scenario of why posting money would be good, and therefore no reason to post money according to REG.<sup>72</sup>

Does Ingrid have a reason to post money in this case? Different considerations push in opposite directions. On one hand, if an advisor like Hannah were to discover this additional information, it would be misleading or deceptive for her to continue telling Ingrid she has a reason to post money, as REG predicts. But on the other hand it seems natural for us to say, in a third-personal context, that Ingrid has a reason to post money. If she were actually to post money we would say she did it for a reason. And if we were asked what that reason was, we might reply that her reason is (roughly) FA, that N is her brother's address.<sup>73</sup> Here we explain her action by citing a "reason" that both motivates and rationally justifies it. This poses a problem for theories like REG that appear to imply there is no reason to act in such cases. A similar problem is posed by cases of *perverse* reasons, where the fact that the agent believes is apparently not a normative (or a "good") reason. For example: 'His reason for pulling his sister's hair was that it would annoy her.'<sup>74</sup> Here too it seems there couldn't be an explanation why acting would be good because that explanandum is false.

A common way of resolving these tensions is to distinguish between *objective* reasons, understood roughly as facts that weigh in favor of acting, and *subjective* reasons, understood roughly as considerations that appear from an agent's perspective to weigh in favor of acting—or alternatively, that rationalize acting.<sup>75</sup> This allows us to disambiguate and say that Ingrid has only a subjective reason to post money, but no objective reason. While REG accounts for objective reasons, it may seem poorly suited to account for subjective reasons due to the explanandum-factivity of explanations. (Explanans-factivity is also a problem, since theories of subjective reasons identify these "considerations" as the *contents* of the agents' beliefs which needn't be factual. To keep these issues distinct, in this section I restrict attention to "subjective" reasons that are factual or the contents of true beliefs, addressing

---

<sup>72</sup> Jacob Nebel (fc) offers an easy escape: he radically denies that normative explanations are explanandum-factive, on the basis of ordinary practices of juxtaposing reasons one ought with reasons one oughtn't. I prefer to interpret these claims compatibly with explanandum-factivity as involving different *ceteris paribus* oughts (i.e. context-shifts).

<sup>73</sup> It seems more natural to identify her reason as *that N is her brother's address and he is thinking of relocating to her city but is short of money*. In our context this helps to clarify Ingrid's relevant information.

<sup>74</sup> Some philosophers might suggest this is merely a "motivating reason", which isn't any kind of normative reason. I address these in section 7. However, subjective reasons can be attributed even when there is no motivation. An agent can be unmoved by her subjective reasons, or even (on some interpretations) unaware that they are reasons for her; for discussion see Finlay 2014: 104-5 and below.

<sup>75</sup> E.g. Schroeder 2007: 14, Way 2009: 4, Parfit 2011a: 35f, cf. Mantel 2014: 54.

potentially non-factive reasons in section 7.) Accommodating this distinction is a challenge for a range of theories of reasons, and so many philosophers conclude there is a fundamental ambiguity in talk about normative reasons.<sup>76</sup> Perhaps one sense can be defined derivatively on the other: “subjective” reasons, for example, as the things an agent believes that would be “objective” reasons if the agent’s beliefs were all true.<sup>77</sup> So defenders of REG could simply scale back their ambitions, and restrict their claims to objective reasons. However, I’ll argue that when combined with ERT, REG needn’t limit its ambitions in this way.

One possible approach is to adopt an “inverted-commas” analysis of these claims. Ingrid’s reason for posting money isn’t really a reason at all, because it is merely Ingrid’s “reason”; i.e. what Ingrid herself takes to be a reason or explanation why posting money would be good, a reason-according-to-Ingrid.<sup>78</sup> This move has some virtues: it saves us from needing to recognize a semantic ambiguity in the word ‘reason’ itself (assuming that inverted commas make a semantic contribution to a sentence), it preserves the explanandum-factivity of genuine reasons-statements while accounting for apparent counterexamples like *Posting Money—In Vain* as involving merely supposed explanations, and extends easily to cases of perverse reasons, which are plausibly reasons-according-to-the-agent.

This solution is arguably not without cost, however. First, it seems revisionary, since reasons statements of these kinds often seem sincere and do not wear scare-quotes on their sleeves. Second, it involves an arguably schizophrenic stance attributing rationality or reasonableness to an agent’s behavior while maintaining the absence of any genuine normative reasons for that behavior.<sup>79</sup> Finally, it may fail to account for all the data: “subjective reasons” statements seem appropriate even in cases where the agent fails to take the consideration to be a reason. Suppose Ingrid fails to recognize FA, the fact that N is her brother’s address, as a reason to post money. Then we could say that FA is a reason that Ingrid has to post money as a way of offering rational criticism of her, but presumably could not mean to say that FA is a reason-according-to-Ingrid.

As I’ll next demonstrate, combining REG with ERT provides the resources for a different solution that unifies “objective” and “subjective” reasons with a single unifying theory, and without any of these potential costs.

### 6.1 Subjective Reasons: A Solution

Observe first that once combined with ERT, REG provides a straightforward solution to the perverse reasons puzzle. In cases of perverse reasons (e.g. *that pulling her hair would annoy one’s sister*), the action might not be good relative to any ends that *we* value, but is good relative to ends desired by the *agent*. The end-relativity of ERT allows us to identify a sense in which the action is good, satisfying the requirement of explanandum-factivity. Neither is this solution ad hoc: in the context of talk about “S’s

---

<sup>76</sup> E.g. Schroeder 2008, Skorupski 2010, Hornsby 2011.

<sup>77</sup> E.g. Schroeder 2009b. This approach struggles to accommodate perverse subjective reasons, unless one opts (as Schroeder does) for a controversial Humean theory on which these would be objective reasons.

<sup>78</sup> Finlay 2014: 109f, Hawthorne & Magidor *fc*.

<sup>79</sup> See Gibbons 2010, Lord 2010, Comesaña & McGrath 2014.

reasons” or “reasons S has”, we should expect that S’s desired ends will be particularly salient, making this a natural reading.<sup>80</sup>

A directly parallel solution is available for the subjective reasons puzzle in virtue of the information-relativity of REG+ERT. As readers may already have noticed, the presentation of the puzzle above assumed an “objective”, or more accurately, speaker or judge-relative sense of value. Ingrid’s posting money doesn’t raise the probability of being reunited with her brother relative to *our* information, once we’ve learned about his impending relationship, and so it isn’t a good thing to do relative to our information (or objectively, relative to all the facts). But posting money does raise the subjective probability of that end for Ingrid, or relative to Ingrid’s information, as we saw in previous sections. By ERT, this means that posting money is subjectively good, relative to Ingrid’s information.<sup>81</sup>

So REG+ERT predicts that there is a sense of goodness on which “subjective reasons” like Ingrid’s in *Posting Money—In Vain* meet the explanandum-factivity condition. The “reason” in this case, (FA) that N is her brother’s address, is indeed an explanation why posting money is good in this sense. So REG+ERT identifies a reading on which FA is a reason for Ingrid to post money. This is also a reading relevant to normative assessments of the *rationality* of the action, which are based not on the facts or what is known to a third party, but rather on the information available to the agent herself. Further, this reading will be especially salient when talking about a reason “for Ingrid”/ that Ingrid “has”, given that REG understands this as an explanation for Ingrid/ that Ingrid has.<sup>82</sup> Relativization or attribution of an explanation to an agent can be expected to direct our attention to that agent’s information, and thereby to what is subjectively good relative to this information. These cues aren’t similarly available when we instead talk about whether it would be good if the agent acts. This explains why the default interpretation of ‘Posting money to N is good’ is of goodness relative to our fuller information (judge-relative or objective goodness), while the default interpretation of ‘Ingrid has a reason to post money’ is agent-relative or subjective, giving rise to the appearance of an explanandum-nonfactivity problem for REG.<sup>83</sup>

The practice of ascribing “subjective” reasons therefore poses no explanandum-nonfactivity problem for REG when it is combined with ERT. On the contrary, REG+ERT is a unifying account that correctly predicts that talk about “reasons” will have both judge-relative and agent-relative (or subjective and objective) interpretations, and when each will be more natural.

---

<sup>80</sup> See Finlay 2006, 2014.

<sup>81</sup> Henning (2014) offers a similar contextualist, information-relative account unifying objective and subjective reasons which I largely endorse (cf. Finlay 2012, 2014, Wedgwood 2015, Hawthorne & Magidor *fc*). But he draws the conclusion that REG is false (p. 621) because he assumes value is always objective, unlike reasons and oughts.

<sup>82</sup> Cf. Henning 2014: 610, 618.

<sup>83</sup> Contrast ‘*There is a reason to post money to N*’, which encourages a more objective/ speaker-relative reading.

## 7. Attitudes as Reasons

Our fifth and final puzzle arises from cases where it seems natural to say that an agent has a reason for acting although the content of the agent's relevant belief, or what they take to be their reason, is *false*. Consider:

*Posting Money—Error:* As in *Posting Money-Explained*, except that Hannah and Ingrid are both misinformed, because N is actually the address of a perfect stranger.

There is still plausibly a sense in which Ingrid has a reason to post money to N in this scenario. In support of this intuition, observe first that given what she believes, posting money would be reasonable, and it would arguably be strange to say an agent acted reasonably although she had no reason for acting that way.<sup>84</sup> Second, if she does post money, it seems correct to say that she does it for a reason, and strange to say she does it for “no reason”. These points are familiar from the puzzle of subjective reasons in section 6. But here we're concerned not with the truth of the explanandum, that posting money to N would be good, rather with the truth of the explanans. The thing that Ingrid believes—the obvious candidate for her reason to post money, (FA) *that N is her brother's address*—is not true.

Many people have the intuition that it is somehow incorrect or infelicitous for us, knowing this, to say that the reason Ingrid has to post money is *that FA*, as this would seem to commit us to the truth of FA. This intuition is not in itself problematic for REG, which naturally accounts for it by the apparent explanans-activity of our concept of explanation. Saying that the explanation of *p* is *that q* seems to imply that *q* is true.<sup>85</sup> The problem rather emerges from what we naturally say about such cases instead. If asked what Ingrid's reason for posting money is, a natural reply would be:

'The reason Ingrid has to post money to N is that *she believes* that N is her brother's address.'

Similarly, if she goes ahead and posts money it seems natural to say,

'The reason for which Ingrid posted money to N was that *she believed* that N is her brother's address.'

Here we apparently identify Ingrid's “reason” as the fact,

IB: Ingrid believes that N is her brother's address.

These responses cite actual (psychological) facts, and as such satisfy the requirement of explanans-activity—but not in a way that looks amenable to REG. For IB doesn't seem like it could be an explanation why posting money would be good or raise the probability of their reunification.<sup>86</sup> This puzzle is again a general one, since IB seems a bad candidate for a normative reason for Ingrid to post money on many theories. As Mark Schroeder writes,

---

<sup>84</sup> Schroeder 2008, Gibbons 2010, Henning 2014, Comesaña & McGrath 2014.

<sup>85</sup> It is, however, a problem for many non-factive analyses of subjective reasons, like Schroeder 2009b, Comesaña & McGrath 2014.

<sup>86</sup> Except perhaps against the background assumption that *Ingrid's beliefs are true*, which doesn't apply here since we cite the agent's belief in case we suspect this belief is false.

Suppose that a fully informed and beneficent bystander is tallying pros and cons of [S's acting]. He's just noted the fact [that *q*] in the cons column. Does he now reflect, "but on the other hand, at least there's this much to say for it—[S *believes* that *p*]?" This seems like a strange thing to say. [S's acting] is no better of an idea, just because he is in the dark... When he is deliberating about what to do, what he believes matters to him only if it is an indication of how things actually are." (2008: 62)<sup>87</sup>

Attentive readers may notice that this presentation of the puzzle seems again to assume a judge-relative or objective sense of goodness or probability. To tip my hand, I will propose below (in section 7.3) a solution in terms of agent-relative or subjective goodness.<sup>88</sup> Our path won't be direct, however, because there are at least two prominent responses to this puzzle already available in the literature. To motivate this new solution I'll first explore why we mightn't be completely satisfied with these alternatives. However, I won't advance it as a *competing* solution, as REG+ERT is sufficiently flexible to be compatible with other readings. A central difficulty in addressing these puzzles is that people's intuitions here are often conflicting or indeterminate.<sup>89</sup> It is another advantage of REG+ERT's flexibility that it can explain and accommodate these divergences by showing how many different legitimate readings of reasons sentences there are.

### 7.1 Attitudes as Reasons: Limitations of the Ambiguity Solution

Perhaps the orthodox response to this puzzle is to disambiguate, following Michael Smith (1994), and to say that reasons-statements that cite the agent's attitudes identify only an agent's *motivating reasons*, which aren't any kind of normative reason at all, but merely an explanation why the agent acted or was motivated.<sup>90</sup> If this *ambiguity solution* is right then there is actually no problem for REG here, since REG is only a theory of normative reasons.

There is much to be said for this response. We do need to recognize a distinction between normative reasons and (nonnormatively) explanatory reasons, which is fully compatible with the broader identification of "reasons" with explanations that inspires REG. Facts about agents' beliefs like IB are often and appropriately cited simply as nonnormative (e.g. causal) explanations or "reasons" why they acted. This seems the correct interpretation of

'The reason why Ingrid posts money to N is that she believes that N is her brother's address',  
which seems functionally equivalent to the nonnormatively explanatory statement,

'The reason why the leaves turned orange was that autumn had arrived'.

Problems potentially arise, however, with reasons statements of the following forms:

---

<sup>87</sup> See also Gardner & Macklem 2004, Alvarez 2010: 139.

<sup>88</sup> Schroeder now also accepts an information-relative diagnosis of this data (ms), although he rejects REG, holding that reasons are normatively fundamental.

<sup>89</sup> For example, Dancy (2000, 2014) holds that a falsehood like FA can be correctly identified as an agent's reason, while Alvarez (2010: 141f) and Hawthorne & Magidor (fc) hold that the agent really does have no reason to act in such cases.

<sup>90</sup> Also Lord 2010. Strictly, Smith identifies motivating reasons not as facts about attitudes, but as the attitudes themselves as the causes of action/motivation. For discussion see Dancy 2000: 122-3, Alvarez 2010: 125f. For a different ambiguity theory appealing to polysemy, see Fogal 2018.

- (3a) The reason for which Ingrid posts money to N is that she believes N was her brother's address;
- (3b) The reason that motivates Ingrid to post money to N is that she believes N is her brother's address;
- (3c) Ingrid's reason for posting money to N is that she believes N is her brother's address;
- (3d) The reason Ingrid has for posting money to N is that she believes N is her brother's address.

One striking thing about these sentences is that they all employ locutions that are clearly infelicitous for obvious cases of nonnormative explanatory statements.<sup>91</sup> Compare:

- (4a) #The reason for which the leaves turned orange was that autumn had arrived;
- (4b) #The reason that motivated the leaves to turn orange was that autumn had arrived;
- (4c) #The leaves' reason for turning orange was that autumn had arrived;
- (4d) #The reason the leaves had for turning orange was that autumn had arrived.

These various locutions we use to attribute agents' motivating reasons strongly suggest that they are about "reasons" of a kind that agents reflect on and are motivated by consideration of, rather than a kind that merely explain why an agent acts. Plausibly it is *normative* reasons, rather than (non-normative) explanatory reasons, that play this role.

The fact that statements of the form of (3a)-(3d) function to explain why an agent acted is often assumed to entail that these "reasons" themselves must be explanations of the action, but this is mistaken. While such *statements* indeed function, at least often, to explain why the agent acted or was motivated, it doesn't follow that the point of calling something a "reason" in these statements is even partly to describe it as an explanation why the agent acted, as opposed to a reason in some normative sense.<sup>92</sup> Suppose someone says, "The *dream* for which he dropped out of school was to become a movie star," or "The *person* for whom she gave up her seat was her brother." It would be absurd to claim that since these statements are explanations of action, the words 'dream' and 'person' here mean *explanation why*. Citing an agent's normative reasons is a legitimate way of explaining why they acted. Having a normative reason isn't sufficient explanation of action all by itself, of course, but the fact that Ingrid had a normative reason to post money is a sufficient explanation why she did against the background assumptions that she was (i) aware of it, (ii) unaware of any stronger countervailing reason, and (iii) disposed to be motivated by the normative reasons she recognizes.<sup>93</sup> Normative reasons are just the kinds of things agents are motivated by.

In the veridical cases where we accept that what the agent takes as a reason is a fact, it is accordingly normal to explain action by citing the fact believed as the reason, rather than the psychological fact that she believes it:

- (5a) The reason for which Ingrid posts money to N is that N is her brother's address;

---

<sup>91</sup> Hawthorne & Magidor (fc) miss this in interpreting statements of the forms in (3a)-(3c) as mere explanatory reasons. Some people report finding these sentences slightly odd; I offer an explanation for this intuition below.

<sup>92</sup> Dancy (2014) notes this point in criticism of his 2000. Cf. Raz 2011: 26.

<sup>93</sup> E.g. Raz 2011: 28f. If these propositions aren't already in the background, the act of explaining an agent's action by citing a normative reason will put them there via accommodation, since we know these are the conditions under which a normative reason standardly explains action.

- (5b) The reason that motivates Ingrid to post money to N is that N is her brother's address;
- (5c) Ingrid's reason for posting money to N is that N is her brother's address;
- (5d) The reason Ingrid has for posting money to N is that N is her brother's address.

Here it is very plausible and widely accepted that these sentences explain the agent's action or motivation by identifying the normative reason, (recognition of) which motivated her to act. This supports an interpretation of an agent's "motivating reasons" as the *normative reasons she had/recognized*, for which she acted/ was motivated; i.e. 'The reason [for Ingrid to post money to N,] for which Ingrid posted money to N, was that N is her brother's address.'<sup>94</sup>

These considerations undermine the motivation for the ambiguity solution, but aren't fatal to it. Plausibly, whenever an agent acts there is a nonnormative explanation why, in addition to any potential explanation by normative reasons. The practice of reporting (the facts that are) normative reasons in the veridical case is something the ambiguity solution can explain. Just as we can explain an action by saying what the normative reason for it was, so too can we explain an action undertaken for a (factual) normative reason that R by identifying R as the (nonnormative) explanation of the action, against the background assumptions that the agent was factually informed and motivated accordingly.<sup>95</sup> Rejecting the ambiguity solution for a normative interpretation of these statements returns us to our puzzle about how psychological facts like IB could be appropriately reported as agents' reasons when they are apparently not relevantly normative reasons.

In addition to the awkward linguistic data sketched above, the ambiguity solution faces a further problem that we sometimes cite psychological facts like IB as "reasons" in contexts that transparently involve normative explanation (or subjective reasons) and not, or not exclusively, explanation of action or motivation.<sup>96</sup> For example, in *Posting Money—Error* it is presumably *rational* for Ingrid to post money, given her beliefs and preferences, and on this basis we can appropriately say that Ingrid rationally should post money, despite our knowledge that her belief is false. This invites the normative question, "What reason does Ingrid have to post money?", to which it seems strange to reply that she has no reason,<sup>97</sup> or that it is FA, that N is her brother's address—which seems to commit us to its truth. It seems more natural to respond with (3c) or (3d), appealing to Ingrid's beliefs.

Although I have been following the ordinary practice of describing these as "motivating reasons" statements, note that these replies—unlike (3a) and (3b)—seem appropriate even if we don't know whether Ingrid will post money or is motivated to do so, and indeed even if we know that she is *not* motivated in this way.<sup>98</sup> Perhaps she has failed to recognize the content of her belief as a reason to act, or is irrationally unmoved by this recognition. As observed in section 6, these statements could be offered as rational criticism of Ingrid. Presumably, therefore, this isn't to identify either a "motivating"

---

<sup>94</sup> See Mantel 2014 for some objections to this line of reasoning.

<sup>95</sup> This explanation also has the resources to explain the asymmetry between veridical and nonveridical cases, in a similar way to the Gricean story I sketch in section 7.3 below.

<sup>96</sup> One could respond by analyzing rationalizing reasons as a certain kind of psychological explanation of action or motivation (Davidson 1963, Williams 1979, Raz 2011: 27)—or vice versa (Smith 1994: 95).

<sup>97</sup> This might be accommodated with a context-shift from a worse (agent's) to a better (judge's) informational background.

<sup>98</sup> Henning 2014: 597.

reason or an explanation of motivation.<sup>99</sup> I'll next turn to strategies for reconciling the puzzle of attitudes as reasons with a normative interpretation.

## 7.2 Attitudes as Reasons: Limitations of the Appositive Solution

The most prominent rival account of attitudes-as-reasons sentences like (3a)-(3d) suggests that their surface grammar is misleading. According to this interpretation, 's believes' is not correctly understood as contributing a component of the normative reason itself, but rather *appositively* (in Jonathan Dancy's terminology) as a parenthetical aside, with the dual functions at least of (i) attributing belief in the supposed reason to the agent, and (ii) distancing the speaker from that belief.<sup>100</sup> On this view, these statements are more transparently formulated as follows:

- (6a) The reason for which Ingrid posts money to N is that, *as she believes*, N is her brother's address;
- (6c) Ingrid's reason for posting money to N is that, *as she believes*, N is her brother's address.

This *appositive solution* also has many virtues. It unifies the "reasons" reported in veridical and nonveridical cases as the same kind of thing, and provides an explanation why we refer to the agent's belief in just those cases where we don't share that belief—i.e. to signal that we don't. It also enjoys a number of advantages over the ambiguity solution.<sup>101</sup> It recognizes that "motivating reasons" statements explain actions by citing what agents took to be their normative reasons, accounts for the difference in locutions between motivating and purely explanatory reasons statements, and extends to cases where an agent's beliefs are apparently cited as "reasons" in contexts of rational criticism and not explanation of action.

Appositivism solves one part of REG's problem with these sentences, since it analyzes away the problematic cases in which facts about attitudes seem to be identified as reasons despite apparently not being explanations why acting would be good. However, in doing so it restores the explanans-factivity problem. According to this view, Ingrid's reason in *Posting Money—Error* is actually FA: that N is her brother's address. But FA is false, so there is no such fact. Mostly, discussion has focused on the question of how these statements could be explanations of action or motivation when they appeal to nonfacts. I don't think this is a serious problem, and endorse Dancy's answer: these motivating reasons statements explain an action by indicating what there was to be said for it from the agent's point of view. Since the appeal to an agent's beliefs creates an intensional context, we can explain this motivation by citing nonfacts.<sup>102</sup> The more relevant problem for our purposes is: how could FA be Ingrid's normative reason, when FA is not the case? If normative reasons are facts, as is widely agreed,

---

<sup>99</sup> Unless perhaps a "motivating reason" is understood as something with the *potential* to motivate; cf. Smith 1994: 96.

<sup>100</sup> Collins 1997, Dancy 2000, Finlay 2012, 2014, Henning 2014. Jacob Nebel suggests (conversation) that 'as she believes, *p*' actually expresses the speaker's *commitment* to the truth of *p*. I can get either reading depending on prosody (whether the appositive is uttered *sotto voce* or with extra stress), but 'so she believes' seems less ambiguous.

<sup>101</sup> They are often treated (e.g. Dancy 2000: 138, 166f) as in competition, although I suggest are better seen as complementary.

<sup>102</sup> Dancy 2000: 146f, Finlay 2014: 113n; cf. Raz 2011: 30. However, I don't agree that the nonfact FA is the real explanans (Dancy 2000: 132), or that this kind of explanation needs no explanans (Dancy 2014). Rather, these statements *implicate* (factual) explanantia, such as IB. But the "reason" cited in these statements isn't the explanans for *why the agent acted*, but rather the (supposed) explanans for *why acting would be good*.

then there is no such reason as FA. How could Ingrid have a reason that doesn't exist? Or how could something be Ingrid's reason if it isn't a reason, "a reason that is no reason"?<sup>103</sup>

Should we reconsider the assumption that explanations are explanans-factive? Dancy suggests we should reject it for explanations (e.g. of motivation or action) involving intensional contexts. However, we saw that on appositivism the point of identifying something as a (motivating) "reason" is not to identify it as an explanation of motivation or action, even if the point of the statement itself is to explain motivation or action. This problem is especially sharp for REG. If we understand reasons as explanations of *goodness*, rather than of motivation, then we can't obviously appeal to the intensionality of belief to account for Ingrid's having a reason.<sup>104</sup>

Might we reject explanans- or reasons-factivity anyway?<sup>105</sup> More can be said for this option than we can discuss here, but it brings further difficulties. What kind of thing is Ingrid's reason on this view? One suggestion is that these are reasons or facts that "exist" but do not "obtain".<sup>106</sup> We may worry about the coherence of this: to call something a "fact" is in part to say it obtains. And if something nonfactual could genuinely be a normative reason, then we might expect to be able (bizarrely!) to cite propositions we know to be false, like FA, in answer to a question, "What reasons are there to act?" (E.g. "Why ought I to do A?" "Because I'll give you a million dollars if you do!" "Will you really?" "No, but that's a pretty good reason to do A!") I'll rather proceed on the assumption that *p* cannot be a genuine explanation why doing A is good if *p* is false, and show instead that there is a solution on REG that doesn't require giving up explanans-factivity.

Another appositivist-friendly suggestion is that we have two different senses of normative reasons here: objective reasons, which are factual, and subjective reasons, which as contents of agents' beliefs needn't be factual.<sup>107</sup> This apparently escapes the explanans-factivity problem by rejecting REG or any explanatory analysis for subjective reasons. It also potentially has costs. First, it multiplies senses of 'reason' in a way we might wish to avoid. Second, if there is such a nonfactive sense of 'reason', then we still face the problem of explaining why falsehoods can't be cited in response to the question, "What reasons are there to act"?<sup>108</sup> Third, if we are using a nonfactive sense of 'reason' then we arguably lack a compelling explanation of why we feel a need to insert 'S believes' in reporting reasons in nonveridical cases. Dancy claims that this indeed isn't necessary, noting the suggestion of factivity can be canceled

---

<sup>103</sup> Dancy 2000: 142.

<sup>104</sup> This issue is obscure in Dancy's treatment, because he doesn't ultimately accept the analysis of reasons as explanations (although he comes close to endorsing REG in 2000: 7-8), rejecting the possibility of an informative analysis. For him, rejecting reasons-factivity therefore doesn't entail rejecting explanans-factivity.

<sup>105</sup> E.g. Dancy 2000: 131f, 146-7, Fantl & McGrath 2009: 100f, Henning 2014: 34, Comesaña & McGrath 2014. Henning suggests that contextualism allows for explanans-nonfactivity, since reasons-claims can be made relative to false information. I am unconvinced, because the reason itself isn't part of the information-base.

<sup>106</sup> Dancy 2000: 146, Comesaña & McGrath 2014.

<sup>107</sup> E.g. Schroeder 2008. Despite originally claiming that "normative" and "motivating" reasons statements apply one concept of a reason to different questions, Dancy concedes (according to Schroeder 2008: 67n) the need to discriminate different senses to make sense of the existence of motivating reasons in the absence of normative reasons, holding that any motivating reason must merely be *capable* of being a normative reason.

<sup>108</sup> Schroeder's (2008) explanation is that subjective reasons are essentially relational: something is a subjective reason *for some agent S*, not tout court. The principle is questionable: parenthood is also an essentially relational property, yet we can intelligibly say which people are parents.

either explicitly or contextually. Even so, however, we need a satisfying explanation for the intuition and tendency.<sup>109</sup>

One further move dissolves all the difficulties just sketched. This is to explain away reasons statements in nonveridical cases as inverted-commas uses of ‘reason’, as described in section 6. Ingrid’s reason for posting money in *Posting Money—Error* isn’t really a normative reason at all, because it is merely Ingrid’s “reason”, or a reason-*according-to-Ingrid* of why posting money would be good. This move doesn’t really multiply senses of ‘reason’, as it rather just observes a familiar, general kind of nonliteral use. It is therefore fully compatible with REG as a unifying analysis, and preserves the explanans-factivity of all *genuine* reasons-statements, explaining away apparent counterexamples like (3a)-(3d) on the appositive reading as involving merely supposed reasons. Since these statements don’t identify what speakers themselves consider genuine reasons, it straightforwardly explains why we wouldn’t cite them when asked what reasons there are (this also extends to the problem of perverse reasons). The use of ‘S believes’ in nonveridical cases is explained by the desideratum of making clear that we as speakers don’t ourselves accept these as facts or reasons. It also explains why the need for ‘S believes’ is defeasible in the way Dancy claims, as the insertion is redundant when speaker nonacceptance is explicit or salient.

Perhaps this is a satisfactory solution to the problem of attitudes-as-reasons, as I have accepted in previous work. However, the residual issues sketched in section 6 might be thought to provide objections. To apply those to our present puzzle: (i) it is an arguably revisionary and nontransparent interpretation of these sentences, since it requires both the controversial appositivist reading of ‘S believes’ and the controversial inverted-commas reading of ‘reason’. (ii) Because it does not recognize the existence of any genuine normative reasons for the action, it fails to account for the agent’s rationality or reasonableness in performing that action. (iii) Relatedly, it may fail to cover all the relevant cases. As observed in section 7.1, attitudes-as-reasons sentences of the forms,

(3c) Ingrid’s reason for posting money to N is that she believes N is her brother’s address;

(3d) The reason Ingrid has for posting money to N is that she believes N is her brother’s address;

can be used not only as “motivating reasons” statements, but as “subjective reasons” statements in cases of rational criticism where the speaker knows that Ingrid is *not* motivated to post money and perhaps doesn’t even recognize her belief FA as a reason to do so. Presumably what we mean by ‘Ingrid’s reason’ or ‘the reason Ingrid has’ in these cases cannot be a *reason-according-to-Ingrid*. We seem rather to be identifying something as having a genuine normative status for Ingrid.

### 7.3 Attitudes as Reasons: An Information-Relativity Solution

Perhaps these concerns are overblown or can be overcome. But we might now observe that REG+ERT provides the resources for a new analysis of such statements that avoids all the problems surveyed above.<sup>110</sup> This solution again turns on the information-relativity introduced by ERT. The fact (IB) that Ingrid believes that N is her brother’s address may not explain why posting money would raise the

---

<sup>109</sup> Dancy 2000: 132-3; see also Comesaña & McGrath 2014 for a pragmatic solution.

<sup>110</sup> This section develops suggestions in Finlay 2014: 113-4.

probability of their reuniting relative to *our* information, and thereby its judge-relative (or objective) goodness.<sup>111</sup> But it *does* in a certain sense explain why posting money would raise the probability of their reuniting relative to *Ingrid's* information, and thereby its subjective goodness for Ingrid.

This is subtly but significantly different from the way we found subjective goodness explained (in section 5) by evidence of objective goodness, like testimony from better-informed advisors, and (in section 6) by incomplete information. The fact that *Hannah says posting money would be good* (HS), in the original *Posting Money* scenario, and the fact that *N is Ingrid's brother's address* (FA), in *Posting Money—In Vain*, are each explanations why posting money would be good relative to what Ingrid believes *de re*. That is to say, they are explanations relative to the extension  $b_i(w)$ , the set of (other) propositions that Ingrid believes. By contrast, in *Posting Money—Error* the fact that *Ingrid believes that N is her brother's address* (IB) is an explanation rather of why posting money would be good relative to what Ingrid believes *de dicto*. That is to say, they are explanations relative to the intension  $b_i$ , *what Ingrid believes*.<sup>112</sup> Against the background information that Ingrid believes that her brother is looking to relocate to her city but is short on funds, IB is an explanation why the probability of their reuniting is higher, relative to Ingrid's beliefs, if she posts money than otherwise: an explanation of *de dicto subjective goodness*.

So REG+ERT predicts a sense of goodness on which facts about agents' reasons-beliefs are indeed explanations of the goodness of acting. This interpretation of "motivating reasons" statements in terms of *de dicto* subjective goodness has many virtues. (i) It preserves the double factivity of reasons/explanations: both the explanandum (the subjective *de dicto* goodness of acting) and—unlike on the appositive reading—the explanans (here, the psychological fact) are factual. (ii) It takes the sentences at face value, and doesn't have to re-parse the "reason" clause in a nontransparent appositive way, or postulate an inverted-commas use of 'reason'. (iii) Because the explanandum is both normative and factual, this analysis can explain (unlike either the nonfactualist or inverted-commas interpretations) why the agent stands in a genuine, and not merely apparent, normative relation to the action. (iv) Furthermore, the kind of normativity involved—involving the pressure that an agent's own attitudes or subjective probabilities place on her—is exactly the kind we should expect to find in such cases; i.e. the normativity of *rationality*.<sup>113</sup> (v) Because this explanandum is normative (why would it be good to act?) rather than behavioral or psychological (why is she motivated to act?), it extends easily to the uses of sentences like (3c) and (3d) in contexts where the agent isn't motivated, such as rational criticism.

---

<sup>111</sup> Dancy dismisses such a solution (2000: 60f), on the grounds that there is only one normative question about what the agent ought to do. I disagree; we can distinguish, for example, between the questions of what it is rational for an agent to do (i.e. given her information), and what it is advisable for that agent to do (i.e. given our information). This is compatible with granting there is only ever one practical or deliberative question about what to do given a particular context of information and goals (Finlay 2014: 154f).

<sup>112</sup> Conversely, FA and HS would not be explanations for you of why posting money to N would be good relative to Ingrid's information *de dicto*, if you weren't aware of the content of Ingrid's information.

<sup>113</sup> The normativity of rationality is controversial (for skepticism about the guiding role of rational requirements see Kolodny 2005; for skepticism about their violability see Finlay 2009, Lee 2018). Here I use 'normative' in a broader sense including the "merely evaluative".

This account is not ad hoc, and easily explains why we asymmetrically switch to reporting this kind of psychological reason only in nonveridical cases. De dicto subjective (or psychological) reasons statements provide less information than de re subjective (or “content”) reasons statements, because the latter convey not only the information about the “actual” reason to act that the agent correctly recognized, but also (implicitly, by accommodation) information about the agent’s attitude and the consequent rational appropriateness of the act, and does so in a particularly succinct way.<sup>114</sup> By Grice’s pragmatic maxims of quantity (“Say the strongest thing you can”) and manner (“Be succinct”), we can therefore expect speakers to report de re or content reasons whenever possible. However, if talk about reasons or explanations is indeed explanans-factive then we cannot felicitously cite these “reasons” when we don’t accept them as true, inverted-commas uses aside. In these cases we can be expected to say the strongest thing that is true, which is the de dicto, psychological reason statement. This statement explains why the agent acted—or, in cases of rational criticism, should have acted—in light of what made it rational to act. Finally, we can explain why we naturally think that facts like IB could not be explanations why acting would be good, creating the appearance of a puzzle for REG. When asked to consider whether whether acting would be good, it is natural to assume the default context where our own fuller information is salient,<sup>115</sup> rather than the context of the agent’s information (and ends) relative to which the action is relevantly good.

I see one potentially significant problem for this interpretation, which arises from sentences like (3a) (“The reason *for which* Ingrid posted money was that she believed...”) and (3b) (“The reason *that motivated* Ingrid to post money...”). Unlike (3c) and (3d), these sentences are unambiguously concerned with motivation rather than what it is rational for the agent to do. But agents aren’t normally motivated to act by reflecting on psychological facts about themselves, or by taking such facts to be reasons—although such higher-order deliberation sometimes happens.<sup>116</sup> If Ingrid were asked, in *Posting Money—Error*, “What was your reason for posting money to N?”, we would expect her to answer “It was that N is my brother’s address”, and *not* “It was that I believed that N is my brother’s address”—prior to discovering her error, at least. So this interpretation suggests that such motivating reasons statements should usually be false, because they misidentify the thing that the agent took to be her reason.

This may ultimately emerge as a virtue rather than a vice of the solution, however. Intuitions and judgments about the felicity of (3a)-(3d) are contested, and we are now in a good position to explain why. Plausibly, it is this problem that underlies resistance to these sentences. Given the close connections and similarities between well-formed reasons statements of the various kinds observed in this section, i.e. *nonnormative explanations* like

‘The reason why Ingrid posts money to N is that she believes N is her brother’s address,’

*supposed normative reasons* statements like

‘The “reason” for which Ingrid posted money to N is that, as she believed, N was her brother’s address,’

---

<sup>114</sup> Cf. Raz 2011: 30f.

<sup>115</sup> Or a strongly objective context: recall Schroeder’s invitation to imagine “a fully informed and beneficent bystander.”

<sup>116</sup> Dancy 2000: 124-5, Schroeder 2008: 62.

and *de dicto subjective normative reasons* statements like

‘The reason Ingrid has to post money to N is that she believes that N is her brother’s address,’

it is hardly surprising if speakers get confused and use an inappropriate or misleading form, or if different theorists fix on different readings of the same sentences, leading to conflicting intuitions. So while the new interpretation proposed here on the basis of REG+ERT cannot explain all the uses and intuitions involving so-called motivating reasons sentences by itself, I suggest it is a necessary part of a complete explanatory package—perhaps even the only piece we were missing.

### Conclusion

Despite its many virtues, the unifying theory that normative reasons are explanations of goodness (REG) faces some *prima facie* serious objections, including the five puzzles explored in this paper. I have argued that solutions to all five puzzles are systematically generated by combining REG with a particular, end-relational theory of goodness (ERT), which wasn’t designed on an *ad hoc* basis for the purpose of solving these puzzles, but rather taken “off the shelf” as an analysis developed in response to independent metaethical and semantic considerations. This constitutes a powerful case for both REG and ERT, and especially their combination, with wide-reaching implications for our understanding of both reasons and value.

There is an important lesson for philosophers writing about reasons, who generally assume it is a simple, yes-or-no question whether a particular fact R is a reason for an agent S to do A.<sup>117</sup> The context-sensitivity of our talk and thought about both goodness and explanation imply rather that the correctness of identifying R as “a reason” for S to do A is heavily dependent on conversational context such as salient ends and background information. Overlooking this makes inevitable the widespread confusion and intractably conflicting intuitions that we find in the contemporary debate over reasons. The good news is that by recognizing and attending to this context-sensitivity these conflicting intuitions can be both explained and reconciled.<sup>118</sup>

---

<sup>117</sup> This point isn’t original to this paper; see e.g. Henning 2014, Finlay 2006, 2014: ch. 4.

<sup>118</sup> Thanks to: Yufan Bi, David Clark, Daniel Greco, John Hawthorne, Jake Nebel, Jake Ross, Daniel Star, Ralph Wedgwood, members of the Board Certified Epistemologists Facebook group, participants of conferences at the Universities of Nottingham (2016), Groningen (2016), Shandong (2018), and Wuhan (2018), and audiences at the philosophy departments at MIT, Boston University, Yale, Princeton, and the University of Connecticut-Storrs. Any errors are mine alone, and I apologize for the many omissions and oversights. Work on this paper was supported by the Office of the Provost and Advancing Scholarship in the Humanities and Social Sciences grant program at the University of Southern California.

## References

- Alvarez, M. (2010). *Kinds of Reasons*. Oxford University Press.
- Behrends, J., & DiPaolo, J. (2011). Finlay and Schroeder on promoting a desire. *Journal of Ethics and Social Philosophy*, 1-7.
- Broome, J. (2004). Reasons. In R. Jay Wallace (Ed.), *Reason and value: Themes from the moral philosophy of Joseph Raz* (pp. 2004-2028). Oxford University Press.
- Broome, J. (2008). Reply to Southwood, Kearns and Star, and Cullity. *Ethics*, 119(1), 96-108.
- Broome, J. (ms). The first normative reason.
- Brunero, J. (2013). Reasons as explanations. *Philosophical Studies*, 165(3), 805-824.
- Chappell, R. Y. (2012). Fittingness: The sole normative primitive. *Philosophical Quarterly*, 62(249), 684-704.
- Collins, A. W. (1997). The psychological reality of reasons. *Ratio*, 10(2), 108-123.
- Comesaña, J. & McGrath, M. (2014). Having false reasons. In Clayton Littlejohn & John Turri (eds.), *Epistemic norms* (p. 59-80). Oxford University Press.
- Cowie, C. (2014). In defence of instrumentalism about epistemic normativity. *Synthese*, 191(16), 4003-4017.
- Crisp, R. (2000). Particularizing particularism. In Brad Hooker & Margaret Little (Eds.), *Moral particularism* (pp. 23-47). Oxford University Press.
- Crisp, R. (2005). Value, reasons and the structure of justification: How to avoid passing the buck. *Analysis*, 65(1), 80-85.
- Dancy, J. (1992). Holism in the theory of reasons. *Cogito*, 6(3), 136-138.
- Dancy, J. (2000). *Practical reality*. Oxford University Press.
- Dancy, J. (2004). *Ethics without principles*. Oxford University Press.
- Dancy, J. (2014). On knowing one's reason. In Clayton Littlejohn & John Turri (eds.), *Epistemic norms* (pp. 81-96). Oxford University Press.
- Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60(23), 685-700.
- Elga, A. (2000). Self-locating belief and the sleeping beauty problem. *Analysis*, 60(266), 143-147.
- Evers, D. (2010). The end-relational theory of 'ought' and the weight of reasons. *Dialectica* 64(3): 405-417.
- Fantl, J., & McGrath, M. (2009). *Knowledge in an uncertain world*. Oxford University Press.
- Finlay, S. (2001). *What does value matter? the interest-relational theory of the semantics and metaphysics of value*. University of Illinois PhD Dissertation.
- Finlay, S. (2006). The reasons that matter. *Australasian Journal of Philosophy*, 84(1), 1-20.
- Finlay, S. (2012). Explaining reasons. *Deutsches Jahrbuch Fuer Philosophie*, 4, 112-126.
- Finlay, S. (2014). *Confusion of tongues: A theory of normative language*. Oxford University Press.
- Fogal, D. (2016). Reasons, reason, and context. In Errol Lord, & Barry Maguire (Eds.), *Weighing reasons*. Oxford University Press.
- Fogal, D. (2018). Deflationary pluralism about motivating reasons. In Veli Mitova (ed.), *The Factive Turn in Epistemology*. Cambridge University Press.
- Fogal, D., & Sylvan, K. (2017). Contextualism about epistemic reasons. In Jonathan Jenkins Ichikawa (Ed.), *The routledge handbook of epistemic contextualism*. Routledge.
- Foot, P. (2001). *Natural goodness*. Oxford University Press.
- Gardner, J. & Macklem, T. (2004). Reasons. In J. Coleman, K. Himma, S. Shapiro (eds.), *Oxford handbook of jurisprudence and philosophy of law*. Oxford University Press.
- Gardner, J. (2008). The logic of excuses and the rationality of emotions. *University of Oxford Legal Research Paper Series* 35.
- Gert, J. (2016). The distinction between justifying and requiring: nothing to fear. In Barry Maguire & Errol Lord (eds.), *Weighing Reasons*. Oxford University Press.

- Gibbons, J. (2010). Things that make things reasonable. *Philosophy and Phenomenological Research*, 81(2), 335-361.
- Gregory, A. (2016). Normative reasons as good bases. *Philosophical Studies*, 173(9), 2291-2310.
- Harman, G., & Thomson, J. J. (1996). *Moral relativism and moral objectivity*. Blackwell.
- Hawthorne, J., & Magidor, O. (fc). Reflections on reasons. In D. Star (Ed.), *The Oxford Handbook on Reasons and Normativity*. Oxford University Press.
- Hedden, B. (2015). Time-slice rationality. *Mind* 124 (494): 449-491.
- Hempel, C. G. (1965). *Aspects of scientific explanation*. Free Press.
- Henning, T. (2014). Normative reasons contextualism. *Philosophy and Phenomenological Research*, 88(3), 593-624.
- Heuer, U. (2006). Explaining reasons: Where does the buck stop? *Journal of Ethics & Social Philosophy*, 1(3), 1-27.
- Heuer, U. (2010a). Beyond wrong reasons: The buck-passing account of value. In Michael Brady (Ed.), *New waves in metaethics*. Palgrave-Macmillan.
- Heuer, U. (2010b). Wrongness and reasons. *Ethical Theory and Moral Practice*, 13(2), 137-152.
- Hieronymi, P. (2005). The wrong kind of reason. *Journal of Philosophy*, 102(9), 437-457.
- Hieronymi, P. (2006). Controlling attitudes. *Pacific Philosophical Quarterly*, 87(1): 45-74.
- Hornsby, J. (2011). A disjunctivist conception of acting for reasons. In Adrian Haddock, & Fiona Macpherson (Eds.), *Disjunctivism: Perception, action, knowledge*. Oxford University Press.
- Howard, C. M. (2017). *The fundamentality of fit*. PhD Dissertation
- Kearns, S., & Star, D. (fc). Weighing explanations. In Andrew Reisner, & Iwao Hirose (Eds.), *Weighing and reasoning: A festschrift for John Broome*. Oxford University Press.
- Kearns, S., & Star, D. (2008). Reasons: Explanations or evidence? *Ethics*, 119(1), 31-56.
- Kearns, S., & Star, D. (2009). Reasons as evidence. *Oxford Studies in Metaethics*, 4, 215-242.
- Kearns, S., & Star, D. (2011). On good advice: A reply to McNaughton and Rawling. *Analysis*, 71(3), 506-508.
- Kearns, S., & Star, D. (2013a). Reasons, facts about evidence, and indirect evidence. *Analytic Philosophy*, 54(2), 237-243.
- Kearns, S., & Star, D. (2013b). Weighing reasons. *Journal of Moral Philosophy*, 10(1), 70-86.
- Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research*, 66(3), 612-640.
- Kiesewetter, B. (fc). How reasons are sensitive to available evidence. In Conor McHugh, Jonathan Way & Daniel Whiting (Eds.), *Normativity: Epistemic and practical*. Oxford University Press.
- Kiesewetter, B. (2011). 'Ought' and the perspective of the agent. *Journal of Ethics and Social Philosophy*, 5(3), 1-24.
- Kolodny, N. (2005). Why be rational? *Mind*, 114(455), 509-563.
- Kratzer, A. (2012). *Modals and conditionals*. Oxford University Press.
- Lee, W. (2018). *Rationality and the Primacy of the Occurrent*. USC PhD dissertation.
- Lord, E. (2010). Having reasons and the factoring account. *Philosophical studies* 149(3): 283-296.
- Maguire, B. (2016). The value-based theory of reasons. *Ergo, an Open Access Journal of Philosophy*, 3.
- Maguire, B. & Snedegar, J. (ms). Explaining overlapping reasons. Manuscript.
- Mantel, S. (2014). No reason for identity: on the relation between motivating and normative reasons. *Philosophical Explorations* 17(1): 49-62.
- McHugh, C., & Way, J. (2016). Fittingness first. *Ethics*, 126(3), 575-606.
- Nagel, T. (1970). *The possibility of altruism*. Oxford Clarendon Press.
- Nagel, T. (1986). *The view from nowhere*. Oxford University Press.
- Nebel, J. (fc). Normative reasons as reasons why we ought. *Mind*.
- Parfit, D. (2011a). *On what matters: Volume one*. Oxford University Press.
- Parfit, D. (2011b). *On what matters: Volume two*. Oxford University Press.
- Philips, M. (1987). Weighing moral reasons. *Mind*, 96(383), 367-375.

- Rabinowicz, W., & Rønnow-Rasmussen, T. (2004). The strike of the demon: On fitting pro-attitudes and value. *Ethics*, 114(3), 391-423.
- Raz, J. (1990) *Practical reason and norms*, 2<sup>nd</sup> ed. Princeton University Press.
- Raz, J. (1999). *Engaging reason: On the theory of value and action*. Oxford University Press.
- Raz, J. (2009). Reasons : Practical and adaptive. In David Sobel, & Steven Wall (Eds.), *Reasons for action*. Cambridge University Press.
- Raz, J. (2011). *From normativity to responsibility*. Oxford University Press.
- Roush, S. (2005). *Tracking truth*. Oxford University Press.
- Scanlon, T. (1998). *What we owe to each other*. Belknap Press of Harvard University Press.
- Schroeder, M. (2007). *Slaves of the passions*. Oxford University Press.
- Schroeder, M. (2008). Having reasons. *Philosophical Studies*, 139(1), 57-71.
- Schroeder, M. (2009a). Buck-passers' negative thesis. *Philosophical Explorations*, 12(3), 341-347.
- Schroeder, M. (2009b). Means-end coherence, stringency, and subjective reasons. *Philosophical Studies*, 143(2), 223-248.
- Schroeder, M. (ms). The fundamental reason for reasons fundamentalism.
- Schroeter, L., & Schroeter, F. (2009). Reasons as right-makers. *Philosophical Explorations*, 12(3), 279-296.
- Searle, J. (2001). *Rationality in action*. MIT Press.
- Setiya, K. (2007). Cognitivism about instrumental reason. *Ethics*, 117(4), 649-673.
- Shah, N., & Velleman, D. (2005). Doxastic deliberation. *Philosophical Review* 114(4): 497-534.
- Shah, N. (2006). A new argument for evidentialism. *Philosophical Quarterly*, 56(225), 481-498.
- Shanklin, R. (2011). *On good and 'good'*. University of Southern California PhD Dissertation.
- Sinnott-Armstrong, W. (2009). How strong is this obligation? An argument for consequentialism from concomitant variation. *Analysis*, 69(3), 438-442.
- Skorupski, J. (2007). Buck-passing about goodness. In T. Rønnow-Rasmussen B. Petersson J. Josefsson D. Egonsson (Ed.), *Hommage `a Wlodek. Philosophical papers dedicated to Wlodek Rabinowicz*.
- Skorupski, J. (2010). *The domain of reasons*. Oxford University Press.
- Smith, M. (1994). *The moral problem*. Blackwell.
- Stratton-Lake, P., & Hooker, B. (2006). Scanlon versus Moore on goodness. In T. Horgan & M. Timmons, *Metaethics after Moore* (pp. 11-21). Oxford University Press.
- Swanson, E. (2010). Lessons from the context-sensitivity of causal talk. *Journal of philosophy* 107(5): 221-242.
- Szabo, Z. G. (2001). Adjectives in context. In Robert M. Harnish, & Istvan Kenesei (Eds.), *Perspectives on semantics, pragmatics, and discourse*. John Benjamins.
- Thomson, J. J. (1992). On some ways in which a thing can be good. *Social Philosophy and Policy*, 9(2), 96.
- Thomson, J. J. (2008). *Normativity*. Open Court.
- Toulmin, S. E. (1950). *An examination of the place of reason in ethics*. Cambridge University Press.
- Väyrynen, P. (2006). Resisting the buck-passing account of value. *Oxford Studies in Metaethics*, 1, 295-324.
- Van Fraassen, B. (1984). Belief and the will. *The Journal of Philosophy*, 81(5), 235-256.
- Way, J. (2009). Two accounts of the normativity of rationality. *Journal of Ethics and Social Philosophy*, 1-8.
- Way, J. (2012). Transmission and the wrong kind of reason. *Ethics*, 122(3), 489-515.
- Way, J. (2013). Value and reasons to favour. *Oxford Studies in Metaethics*, 8.
- Way, J. (2017). Reasons as premises of good reasoning. *Pacific Philosophical Quarterly*, 98(2).
- Wedgwood, R. (2015). The pitfalls of reasons. *Philosophical Issues*, 25(1), 123-143.
- Wedgwood, R. (ms). The reasons aggregation theorem.
- Williams, B. (1973). Deciding to believe. In *Problems of the Self*. Cambridge University Press.
- Williams, B. (1979). Internal and external reasons. In Ross Harrison (Ed.), *Rational action* (pp. 101-113). Cambridge University Press.

Wolfsdorf, D. (fc). *On Goodness*.

Woodward, J. (2014). Scientific explanation. *Stanford Encyclopedia of Philosophy*, Retrieved from <https://plato.stanford.edu/entries/scientific-explanation/>

Ziff, P. (1960). *Semantic analysis*. Cornell University Press.