

# How to Be a Prudential Expressivist

JAMES L. D. BROWN   
University of Sheffield, UK  
[j.ld.brown@sheffield.ac.uk](mailto:j.ld.brown@sheffield.ac.uk)

This paper examines the prospects for an expressivist theory of prudential thought and discussion, or thought and discussion about what is good for us or what makes our lives go well. It is becoming increasingly common to view prudential thought and discussion as a kind of normative thought and discussion. If this is right, then expressivism, like any other meta-normative view, must be able to explain prudential thought and discussion. However, existing expressivist theories offer no such explanation and lack the resources to construct one. I argue that the best strategy for expressivists is to adopt a fitting attitudes account of prudential concepts. More specifically, I propose that expressivists adopt the rational care theory of well-being, according to which claims about what is good for a person are equivalent to claims about what it is rational to want for that person insofar as one cares for them. In doing so, I defend the rational care theory against its most pressing objection and argue that the view provides an independently attractive account of prudential thought and discussion that fits well with the expressivist's aim to explain normative thought and discussion in terms of its distinctive practical function.

## 1. Introduction

Posing and answering questions about how we should act, what feelings we should have, and what we should believe is a central feature of human life. Call this activity *normative* inquiry. In addition to engaging in this kind of activity, we sometimes take a further reflective step back and ask what, exactly, we are doing when we engage in normative inquiry. Call this activity *meta-normative* inquiry. Throughout much of its history, meta-normative inquiry has tended towards a peculiarly *moral* or *ethical* focus—hence its more common name, *meta-ethics*. Although this history displays no consensus about what exactly constitutes morality or ethics—or indeed about whether these terms are synonymous—we nonetheless see a particular concern to explain such things as the nature of *moral* judgment and the existence of *moral* facts. However, it is now generally recognized that many, if not all, of the philosophically puzzling features of morality that spur meta-ethical inquiry are in fact features of normativity more generally. Thus, many theorists have shifted their focus from morality to normativity more broadly, where the latter includes but is not exhausted by the former.

If the normative domain is not exhausted by morality, what else does it include? Much recent meta-normative inquiry has been dedicated to answering this question.<sup>1</sup> A common assumption of these debates is that the normative domain divides relatively naturally into distinctive subdomains. One such putative subdomain that is receiving increasing attention is that of *prudential normativity*. By ‘prudential’, I mean that which relates to well-being, welfare, self-interest, personal good, or what we might ordinarily describe in terms of what benefits or harms us, what makes our lives go well or badly, or what is good or bad for us.<sup>2</sup> Although it is not uncontroversial, it is becoming increasingly common to consider prudential thought and discussion as a component of normative inquiry. However, the meta-normative implications of this view have been underexplored.<sup>3</sup>

For instance, if prudential thought and discussion is a component of normative inquiry, then any mature meta-normative theory must explain this fact. Specifically, it must explain (a) what makes prudential thought and discussion normative and (b) what distinguishes it from other parts of normative inquiry. We might try to provide an explanation of (a) and (b) that is neutral between differing meta-normative theories. However, there is no reason to expect that different kinds of meta-normative theory will converge on answers to such general questions as what makes something normative. Instead, it seems to me that progress is best achieved piecemeal by developing particular theories of prudential thought and discussion that explain (a) and (b) on their own terms. On this approach, a natural starting point is to examine what our best meta-normative theories might say about prudential thought and discussion.

<sup>1</sup> For examples of recent contributions to these debates, this is asked of *structural rationality* (for example, [Worsnip 2021](#)), *epistemic norms* (for example, [Maguire & Woods 2020](#)), *political norms* (for example, [Maynard & Worsnip 2018](#)), *aesthetics* (for example, [Hills forthcoming](#)), *legal norms* (for example, [Plunkett, Shapiro & Toh 2019](#)), and so on. These debates are often couched in terms of whether some domain is *robustly*, *substantively*, or *authoritatively* normative, as it is usually uncontroversial that the domain in question is *formally* or *generically* normative (that is, involves standards that evaluate and rank options using evaluative and deontic vocabulary). Throughout, I use ‘normative’ to denote robust normativity; for an account of what this is, see [Brown \(2024\)](#).

<sup>2</sup> Some might object to the identifications made here. For instance: [Railton \(1989\)](#) distinguishes personal good from well-being; [Kagan \(1994\)](#) distinguishes well-being from a life’s going well; and [Ridge \(2024\)](#) distinguishes prudence from well-being. While I think such views are mistaken, for reasons of space I will simply assume that these terms denote a unified kind. It is worth noting, however, that ‘prudential’ is not an entirely satisfactory label for the relevant kind insofar as ordinary uses of ‘prudence’ more commonly refer to something like the quality of cautiousness or showing care for the future. But there is also an ordinary use of ‘prudence’ that means simply a regard for one’s interests.

<sup>3</sup> Prominent exceptions include [Railton \(1989\)](#); [Darwall \(2002\)](#); [Campbell \(2016\)](#); [Fletcher \(2021\)](#); [Lin \(2022\)](#).

In this spirit, this paper examines the prospects for an *expressivist* theory of prudential thought and discussion. According to expressivism, normative claims express non-representational, practical attitudes (see, for example, Blackburn 1984, 1998; Gibbard 1990, 2003; Schroeder 2008; Ridge 2014; Sinclair 2021; Lenman 2024). On this view, normative thoughts do not track or represent features of reality; rather, they play a conative role in the production of action and the regulation of attitudes. Many consider expressivism an attractive meta-normative view insofar as it provides a straightforward explanation of the practical nature of normative inquiry that vindicates our normative practices without positing a putatively problematic domain of normative facts. Thus, if expressivism provides an attractive meta-normative view, and prudential thought and discussion is normative, then we should expect expressivism to provide an attractive view of prudential thought and discussion. The problem, however, is that extant expressivist theories offer no account of prudential thought and discussion. Moreover, it is not at all obvious that what expressivists do say about normative inquiry can be applied to prudential thought and discussion. The first main part of this paper argues that prudential normativity thus constitutes a *challenge* for expressivism (§2).

The second main part then develops an expressivist account of prudential thought and discussion that answers this challenge (§3). I argue that expressivists should explain the normativity of prudential thought and discussion in terms of prudential claims expressing whatever attitude is expressed by normative claims more generally. The more difficult task for the expressivist is to explain what distinguishes prudential attitudes from other kinds of normative attitudes. In relation to this task, I argue for two claims. First, that fitting attitude accounts of normative concepts provide a promising way of explaining different varieties of normativity within an expressivist framework. Second, that Darwall's (2002) rational care theory of prudential concepts provides an attractive account of prudential thought and discussion that can be utilized by expressivists. The basic idea is that what makes prudential claims distinctive is that they are claims about what one should want for those that one cares about, where 'should' expresses a basic non-representational, practical attitude common to all normative claims.

While my focus here is on prudential normativity, our discussion can be understood as part of a wider investigation into how to make sense of the varieties of normativity more generally and their relation to one another. So if the strategies for explaining prudential normativity work in this case, it will be instructive to see how far they can generalize to other varieties of normativity.

## 2. The challenge from prudential normativity

In this section, I argue that expressivism faces a challenge explaining prudential thought and discussion. However, it is not uncontroversial that prudential thought and discussion is in fact normative.<sup>4</sup> If it is not normative, then expressivists do not face any challenge explaining it. So we first need to say something about why expressivists need to explain prudential thought and discussion at all.

### 2.1 *The normativity of prudential thought and discussion*

To begin, it will be helpful to get clearer on what prudential thought and discussion is. As I am using the term, ‘prudential’ denotes that which concerns a specific kind of value and consideration. The kind of value is *prudential value*, the value something has with respect to an individual’s well-being, welfare, or personal good. In ordinary prudential discussion, this value is denoted by expressions such as ‘good for’, ‘better for’, ‘benefit’, ‘self-interest’, and so on.<sup>5</sup> The kind of considerations are *prudential reasons*, which are in some way distinctively about or explained in terms of the promotion of an agent’s well-being. When we consider prudential reasons in abstraction from other kinds of reasons, they will also determine distinctively prudential interpretations of deontic expressions like ‘should’, ‘ought’, ‘need’, and so on.

The prudential interpretation of evaluative and deontic expressions is typically supplied by the conversational context rather than being explicitly marked as prudential. However, a prudential interpretation can always be forced by an ‘in view of’ phrase, such as ‘in view of your well-being...’ (c.f. [Kratzer 1977](#)). Consider, for instance, the following representative examples of ordinary prudential talk (taken from [Fletcher 2021](#), p. 2):

- [1] ‘You’ll be better off if you have surgery.’
- [2] ‘Her spouse’s death was terribly bad for her.’
- [3] ‘Sarah should take the job.’
- [4] ‘You need friends in order to live well.’

Although other readings of these sentences might be available depending on the context, each sentence allows for a distinctively prudential reading. In (1) and (2), we can interpret the evaluative expressions as denoting the value of things in relation to the well-being of the agent in

<sup>4</sup> See the various citations given by [Fletcher \(2021, pp. 4-5, 33-4\)](#) of those who take it to be obvious that prudential thought and discussion is or is not normative.

<sup>5</sup> For simplicity, throughout I mainly focus on prudential value rather than disvalue.

question. In (3) and (4), we can interpret the deontic expressions prudentially, where the standard or ranking relative to which the sentence is evaluated is determined solely by prudential considerations.

A compelling *prima facie* motivation for the claim that prudential thought and discussion is normative is that it displays several key ‘markers’ of normative inquiry. For instance, prudential judgments (i) are expressed using evaluative and deontic vocabulary, (ii) have authority in deliberation, (iii) are connected with affective and motivational states, and (iv) allow for radical disagreement over the subject matter of prudence (Fletcher 2019). Prudential judgment thus bears a close resemblance to moral judgment, which is often taken as the paradigm of normative judgment. If we were looking for an explanation of why prudential judgments share these features, the simplest would be that prudential judgments just are normative judgments.

Moreover, some of these features are precisely the features of normative inquiry that expressivists believe require an expressivist explanation. For instance, consider the following remarks from Blackburn:

The reason expressivism in ethics has to be correct is that if we supposed that belief, denial, and so on were simply discussions of a way the world is, we would still face the open question. Even if that belief were settled, there would still be issues of what importance to give it, what to do, and all the rest. For we have no conception of a ‘truth condition’ or fact of which mere apprehension by itself determines practical issues. For any fact, there is a question of what to do about it. But evaluative discussion just is discussion about what to do about things. (1998, p. 70)

Like evaluative discussion more generally, prudential discussion also seems to have authority in determining practical issues. As Railton (1989, p. 151) remarks, ‘someone who spoke in earnest to others about their own good, and was simply puzzled when they took his remarks to be any sort of recommendation, would betray a lack of full competence with such discourse.’ So if factual discussion cannot tell us what to do and prudential discussion can, then prudential discussion cannot be factual.

Consider also Köhler’s (2021, p. 618, emphasis in original) claim that ‘one of the *primary* motivations for expressivism’ is the argument from disagreement. Ordinarily, when two speakers apply a single descriptive term according to radically different criteria, we ascribe different contents to the concepts they express with that term. Because of this, such divergences do not constitute genuine disagreements. This contrasts with cases in which two speakers might apply a normative term like

‘ought’ or ‘good’ according to radically different criteria while intelligibly disagreeing over a common subject matter. The best explanation of this, according to expressivists, is that normative disagreements are practical rather than factual disagreements. In much the same way, prudential discussion also seems to allow for radical disagreement. Consider, for instance, a disagreement between a hedonist and an ascetic about whether the best life involves attaining or foregoing pleasure. Or consider the disagreement between Socrates and Polus in the *Gorgias* about whether he is worst-off who does wrong or who suffers it. If such disagreements are possible, and the argument from disagreement is sound, then prudential disagreements are best explained as practical rather than factual disagreements.

Clearly, there is much more to say about these arguments. But given their importance in motivating expressivism about normative thought and discussion, it seems that similar considerations should lead expressivists to extend their view to prudential thought and discussion. Hereafter, I will simply assume that prudential thought and discussion is in fact normative and explore the consequences of this assumption for expressivism. What is clear is that, given this assumption, a mature expressivist theory must be able to account for prudential thought and discussion. In the remainder of this section, I will argue that explaining prudential thought and discussion is not simply unfinished business for expressivism but constitutes a challenge for the view.

## 2.2 *The challenge*

It is important first to distinguish our challenge from a different though related challenge that sometimes goes by the name of the ‘moral attitude problem’. This is the problem of explaining how to distinguish moral or normative judgments from non-moral or non-normative conative attitudes, on the assumption that the former are a kind of conative attitude.<sup>6</sup> By contrast, the challenge from prudential normativity arises at a different stage of theorizing. Specifically, our challenge is *not* to explain what distinguishes normative from non-normative thought. Rather, it is to explain what distinguishes a certain *kind* of normative thought from *other* kinds of normative thought. So the challenge from prudential normativity presupposes an answer to the moral attitude problem. In what

<sup>6</sup> See Smith (2001, 2002); the name comes from Miller (2013). Given that most serious expressivists reject any straightforward identification between normative judgments and basic conative attitudes like desires, preferences, or plans (see below), it is not obvious that the moral attitude problem constitutes a genuine problem for expressivists.

follows, I examine three prominent expressivist theories and argue that none of them contains the resources to explain prudential thought and discussion.

The first theory comes from Blackburn (1998). According to Blackburn, normative claims express complex clusters of conative and affective dispositions. It is not entirely clear how best to frame Blackburn's account in relation to our debate, as Blackburn does not sharply distinguish the moral or ethical from the normative. However, I suggest that we read his account in the following way. Normative claims are those that express *valuing* attitudes, where to hold a value 'is typically to have a relatively stable disposition to conduct practical life and practical discussion in a particular way: it is to be disposed or *set* in that way, and notably to be set against change in this respect' (Blackburn 1998, p. 67, emphasis in original). To be set in this way involves a disposition to resist such change or feel pain when what we value is threatened or taken away. Blackburn then individuates specifically moral attitudes as (roughly) those valuing attitudes that include a disposition to approve or disapprove of those who share or lack that attitude and to approve or disapprove of others' approval or disapproval of having that attitude (1998, p. 9). However, supposing that Blackburn's account successfully allows us to distinguish moral valuing attitudes from non-moral valuing attitudes, there is nothing in his account that allows us to distinguish between other kinds of valuing attitudes. So wherever prudential valuing falls on the ladder of 'emotional ascent', we have not explained what distinguishes prudential thought from other kinds of normative thought.<sup>7</sup>

The second theory comes from Ridge (2014). According to Ridge, normative claims express hybrid attitudes that contain both a representational and a non-representational component. The non-representational component is what Ridge calls a 'normative perspective'. Roughly, this is a kind of practical stance that provides the agent with a set of policies about which standards of practical reasoning to accept and reject (2014, p. 115). The representational component is a belief about how something is ranked according to the standards one accepts and rejects (2014, p. 119). For instance, and very roughly, the claim that 'stealing is wrong' expresses (a) a normative perspective and (b) the belief that stealing is ruled out by (a). While Ridge provides a clear account of which conative attitudes are normative, it is not immediately clear how his account

<sup>7</sup> Miller (2013, pp. 82-3) makes a similar point with respect to aesthetic judgments and judgments of gustatory taste. See also Köhler (2013, p. 486).

can explain the varieties of normativity.<sup>8</sup> Indeed, Ridge seems to use the terms ‘normative’ and ‘moral’ interchangeably (at least, in the context of practical normativity). Given the role that standards play in Ridge’s theory, a tempting answer might be that prudential claims are sensitive to one’s *prudential standards*, understood as the subset of standards not ruled out by one’s normative perspective that relate to prudential considerations. However, the expressivist needs to provide a characterization of these standards in non-normative, and thus non-prudential terms. So the suggestion does not provide an informative explanation of what distinguishes prudential thought from other kinds of normative thought.

The third theory comes from Gibbard (1990, 2003).<sup>9</sup> According to Gibbard, normative claims express plan-like attitudes of norm-acceptance. Very roughly, to accept a system of norms is to plan to act in accordance with what those norms prescribe. Normative discussion is thus akin to contingency planning. It aims to settle how to act, think, and feel in various real and imagined circumstances. While norm acceptance is *plan-like*, we should be careful not to identify normative judgments with plans in the ordinary sense of the term. For one thing, we cannot plan how to feel or what to believe. For another, normative judgments can participate in theoretical reasoning in a way that plans cannot. But, importantly, both kinds of attitude are fundamentally directive in the role they play within our cognitive economy. The precise nature of norm acceptance is explained on this account by specifying the *sui generis* functional role of normative attitudes, namely their distinctive motivational role within an agent’s cognitive economy (what Gibbard calls ‘normative governance’) and their coordinative role within interpersonal normative discussion (1990, p. 75). Gibbard then proposes that specifically moral claims express states of norm-acceptance concerning when to feel guilt and anger (1990, p. 47). We will return to this idea in a little more detail in the next section, as it will provide a useful model for answering the challenge from prudential normativity. All I will note

<sup>8</sup> See Wodak (2017) for a related, though different criticism. Wodak’s worry is that Ridge provides an implausibly disunified meta-semantics for evaluative and deontic expressions, because Ridge provides a different meta-semantic story for normative and non-normative uses of these expressions. My objection pertains solely to Ridge’s meta-semantics for normative uses—specifically, it is that Ridge’s meta-semantics fails to distinguish between different flavours of normative expressions.

<sup>9</sup> This sketch draws on aspects of Gibbard’s earlier and later presentations of his view. While Gibbard uses somewhat different terminology in his (1990) than in his (2003), I do not think this reflects any substantive differences. For an example of Gibbard employing the terminology of both interchangeably, see his (2006).



for now is that Gibbard's theory offers no suggestions regarding how to distinguish prudential thought and discussion from other kinds of normative inquiry.

The lesson here is not that these theories *cannot* accommodate prudential thought and discussion. Indeed, I will argue that the positive proposal suggested in the next section is compatible with all three theories. Rather, the lesson is that these theories *do not* explain prudential thought and discussion as they currently stand. Moreover, filling this explanatory lacuna is not simply a matter of applying resources already contained within these theories to the prudential domain. More needs to be said.

Granting that expressivists have not yet explained prudential thought and discussion, one might wonder to what extent this is really a problem for expressivism. Insofar as all meta-normative theories must explain prudential thought and discussion, isn't prudential normativity *everyone's* problem? It would be beyond the scope of this paper to assess how other theories might answer the challenge. However, it is worth highlighting that there is a simple strategy available to representationalist meta-normative theories that is not available to expressivists. Specifically, representationalist theories can say that prudential thought and discussion is thought and discussion about a particular kind of *fact*: facts about well-being and prudential value. However, given that expressivism is a non-representationalist meta-normative theory, this strategy is not available to it. So the challenge is especially pressing for expressivists.<sup>10</sup>

### 3. Meta-prudential expressivism

In this section, I show how expressivists can answer the challenge of prudential normativity. First, I argue that fitting attitude accounts of normative concepts provide a general way of individuating normative domains within an expressivist framework. Second, I argue that Darwall's rational care theory of prudential concepts provides a promising proposal for meta-prudential expressivism.

#### 3.1 *Fitting attitudes and the varieties of normativity*

Our challenge is to explain (a) what makes prudential thought and discussion normative and (b) what distinguishes it from other kinds

<sup>10</sup> Though this representationalist strategy might face some challenges of its own—compare Maynard & Worsnip (2018, pp. 761-2) and Wedgwood (2007, pp. 18-23).

of normative inquiry. In this sub-section, I argue that *fitting attitude* accounts of normative concepts provide a promising approach to explaining prudential thought and discussion within an expressivist framework. Fitting attitude accounts are typically given as accounts of our *value* concepts, though we will see below that they can also be given for other kinds of normative concepts.<sup>11</sup> For instance, a fitting attitude account of *being good* might identify this concept with that of *being a rational object of desire*, where this might also be expressed in terms of being an *appropriate* or *fitting* object of desire, or being something we *ought* or have *reason* to desire.<sup>12</sup>

Fitting attitude accounts thus explain normative concepts in terms of something non-normative (that is, certain responses) and some more fundamental, primitive normative concept (that is, *rational* or *fitting* or ...). This suggests the following strategy for individuating normative domains within an expressivist framework. First, we identify the set of responses the rationality (or fittingness, or ...) of which is the distinctive subject matter of the domain in question. This explains what distinguishes the domain from other normative domains. Second, we provide an analysis of ‘rationality’ (or ‘fittingness’, or ...) in terms of its expressing whatever basic attitude expressivists take to be constitutive of normative thought. This explains what makes the domain in question normative. Thus, we individuate different kinds of normative attitudes in terms of the distinctive *objects* of a fundamental normative attitude common to all normative thought.

If this strategy seems familiar, it should, because we have already encountered an example of it. In the previous section, we saw how Gibbard explains moral claims in terms of their being about the rationality of guilt and anger—that is, when it is appropriate to feel guilt and anger. Given his account of what is expressed by ‘rational’, to judge that some action is wrong is (roughly) to accept norms that dispose one to feel guilty when one performs that action and to feel anger at others who perform that action. So Gibbard’s account of moral concepts is itself a kind of fitting attitude account.

Importantly, however, an expressivist can accept this account of moral concepts without adopting Gibbard’s norm-expressivism. All

<sup>11</sup> For an overview of this approach and its history, see Rabinowicz (2013). Fitting attitude accounts might also be given as real definitions of normative properties. However, given expressivism’s non-representational commitments, my focus here will be on concepts.

<sup>12</sup> Of course, these formulations are not equivalent. As I am mainly concerned with the basic shape of the account, and as those I discuss use these formulations interchangeably, I will gloss over these details here.

they need is some fundamental normative attitude that is expressed by 'rational' in the analysis of moral concepts. For instance, if normative perspectives can rule out standards for when to feel certain emotions, then Ridge's theory can say that moral claims express hybrid states that involve a normative perspective and a belief that guilt and anger are required or recommended by the standards not ruled out by that perspective.<sup>13</sup> Similarly, Blackburn can say that moral claims express valuing attitudes concerning when to feel guilt and anger. Moreover, because fitting attitude accounts are not peculiar to moral concepts, it is open to expressivists to explain prudential and other kinds of normative concepts in this way. However, the viability of this approach depends on the availability of a plausible fitting attitudes account of our prudential concepts. The remainder of the paper examines one such account: the rational care theory.

### 3.2 Rational care

**3.2.1 The rational care theory** What, then, is the set of responses the rationality of which is the subject matter of prudential thought and discussion? I propose to answer this question by appealing to the rational care theory of well-being, most closely associated with Darwall (2002).<sup>14</sup> According to Darwall, prudential thought and discussion is about the rationality of *care*. At the centre of his account is the proposal that our concept of *being good for x* just is the concept of *being rational to want for x insofar as one cares for x*. Thus, for instance, to claim that a new job would be good for Sophie is to claim that, insofar as one cares for Sophie, it is rational to want that she gets a new job. Similarly, to claim that I would be better off if I had surgery is to claim that it is rational for those who care about me to desire that I have surgery more than the salient alternatives. Thus, prudential claims are claims about what we should want for those we care about (including ourselves).<sup>15</sup>

Importantly, the rational care theory is neutral with respect to which things we should want for those we care about. As such, the theory can be understood as a purely meta-normative account of our prudential

<sup>13</sup> It is important to keep in mind that the sense of 'rational' employed here is not that of structural rationality, which Ridge denies is robustly normative (see Ridge 2014, ch. 8). Rather, the sense of 'rational' relates to what reasons one has or what is fitting. For discussion of this use of 'rational' in an expressivist context, see Gibbard (1990).

<sup>14</sup> See also Anderson (1993) and Rowland (2016). For an alternative fitting attitudes account of prudential concepts, see Rønnow-Rasmussen (2011).

<sup>15</sup> For ease of exposition, I am skipping over some of the nuances of how best to formulate the rational care theory, for which Darwall provides several non-equivalent formulations. For discussion, see Feldman (2006); Darwall (2006); Lin (2022, pp. 6-7).

concepts. Moreover, insofar as the theory offers a fitting attitude account of our prudential concepts, it would seem to offer an attractive and independently motivated way for expressivists to explain prudential thought and discussion. For we can interpret the theory as holding that prudential claims express attitudes of norm acceptance, normative perspectives, valuing attitudes, or whatever, concerning what attitudes to have towards those we care about. Moreover, as I will explain in more detail below, the rational care theory answers the expressivist's desire to explain normative practice in terms of its distinctive practical function—in this case, regulating and coordinating our caring attitudes and behaviours.

However, as has been pointed out by virtually every commentator on the rational care theory, the theory appears to be viciously circular. For we have explained prudence in terms of rational care, but it seems that we cannot explain the nature of care without appealing to prudential notions. For does caring for someone not involve being responsive to their welfare? This is a serious problem for the rational care theory, but even more serious for the expressivist, who aims to provide a non-normative account of normative thought. However, examining the problem will provide us with an opportunity to develop a richer picture of prudential thought. It will also show us how resources developed by expressivists elsewhere can be used to defend the rational care theory. Thus, we will see that expressivism and the rational care theory are mutually reinforcing and together constitute an attractive package, even if they are logically independent.

**3.2.2 *The nature of care*** What is it to care for someone? According to Darwall, when we care for someone, we want things *for that person's sake*, in the sense that we want those things 'with attention to or in consideration of' (2002, p. 68) the person themselves, where this involves 'a whole complex of emotions, sensitivities, and dispositions to attend' (2002, p. 2). However, while Darwall goes to some length to distinguish care from other nearby attitudes, particularly various forms of empathy, he says little about how we should understand the complex of emotions, sensitivities, and dispositions that constitute caring for someone.

A more informative account can be found in Jaworska:

Typical components of caring include: joy and satisfaction when the object of one's care is flourishing and frustration over its misfortunes; anger at agents who heedlessly cause such misfortunes; pride in the successes of the object of care and disappointment over its failures; the desire to help ensure those successes and to help avoid the failures; fear when the object of care is in danger and relief when

it escapes unharmed; grief at the loss of the object, and the subsequent nostalgia. (2007, p. 560)

This characterization requires some amending. For instance, the components of care are more plausibly characterized not directly in terms of how the object of care is *in fact* faring, but rather in terms of how the object *seems* to be faring to the subject. If it seems to me that my beloved is in danger, although they are in fact not, I will still be fearful. This might be because my beliefs about the world are mistaken, but it might also be because I have a mistaken conception of what constitutes misfortune or harm. Thus, the components of caring must be read *de dicto* and not *de re*. Otherwise, however, this seems like a highly plausible gloss on the emotions, sensitivities, and dispositions involved in caring.

Moreover, whereas Darwall's account seems to imply that care is an attitude that we take exclusively towards people, Jaworska's characterization makes clear that the attitude of caring is broader than this. This is important when we consider that the rational care theory is meant to provide an account of well-being that is neutral with respect to first-order questions about well-being. However, restricting the domain of well-being subjects to people seems to determine in advance that only people can be well-being subjects. For this reason, the term 'care' seems preferable to Darwall's other labels of 'sympathy' or 'sympathetic concern'.<sup>16</sup>

In the context of the rational care theory, the problem with Jaworska's characterization is that it makes use of the very prudential concepts the rational care theory sets out to explain. We have explained well-being in terms of care, but then we have explained care in terms of flourishing, misfortune, and so on. Indeed, this is also a feature of Darwall's own positive characterizations of care; for example, in a representative statement of his view, Darwall writes that care 'involves concern for another in light of apparent obstacles to her welfare' (2002, p. 69). Aware of the circularity, Darwall argues that because care is a psychological natural kind, it is not the kind of thing that requires a definition; *a fortiori*, it does not require a definition that mentions well-being. However, if we still need to mention well-being when fixing the reference of 'care', or if we still need to mention well-being when explaining the nature of care, then the account will still exhibit circularity even if we reject the need for a prior definition.

<sup>16</sup> This is supported by the fact that Anderson, who Darwall cites as a progenitor of the rational care view, explicitly takes care to apply to 'people, animals, communities, and things' (1993, p. 20 and *passim*).

In a later discussion of the circularity problem, Darwall (2006, p. 651) concedes that ‘the very idea of caring for someone involves desiring that person’s good.’ In fact, this commitment is already clear in Darwall’s original presentation of his view: ‘What is a conceptual truth is that to care for someone is to be in a relation to him such that considerations of that person’s welfare are normative for one’s desires and actions with respect to him (2002, pp. 6-7, emphasis in original). Given this, a possible response might be to accept the circularity but deny that it is vicious by treating the rational care theory as a non-reductive analysis (Shah 2004, p. 281). After all, such a theory might still be true and informative. The problem for us, however, is that whatever the merits of this response for Darwall, it appears to be incompatible with expressivism. This is because expressivists aim to provide a *non-normative* account of our normative attitudes. And insofar as we appeal to prudential notions to explain care, we have failed to provide such an account. Thus, some other solution is needed.

If care is a complex cluster, then to explain the nature of care, we must explain its components and how they compose to form the cluster. Rather than detailing each component, my strategy will be to show how to provide a non-circular, non-normative account for a single component of care. What I will say about this component can then be applied *mutatis mutandis* to the other components. Given its centrality in discussions of the rational care theory, I will focus on the desire component of care—specifically, desiring that the object of care fares well.

**3.2.3 Solving the circularity problem** What seems to have gone unnoticed in the many discussions of the circularity problem is that the rational care theory *can* explain the nature of care without mentioning well-being. If caring for someone involves desiring that they fare well, this is equivalent to saying that caring for *x* involves desiring that which it is rational to want insofar as one cares for *x*, where as before this is read *de dicto* rather than *de re*. Because this second formulation makes no mention of well-being, we can eliminate any reference to well-being in our characterization of care. While this goes some way to solving the problem, it raises two worries of its own. First, our account apparently fails to be non-normative, since we have explained care partly in terms *rationality*. Second, our account apparently fails to be non-circular, since we have explained care partly in terms of *care*. It will be helpful to separate the two aspects of the challenge. So let’s begin with the normativity worry.

The normativity worry arises because our account partly explains care in terms of something normative, namely rationality. Expressivists, however, eschew explanatory appeals to normative notions within their

metaethical theory. So we need some way of reconciling the appeal to rationality with expressivism's non-representationalist commitments. The solution can be seen when we examine the precise role that 'rational' plays in our account. Specifically, 'rational' is used to individuate a particular attitude type—desiring that which it is *rational* to desire in some context or other (we can bracket any mention of care for the moment). It does this by featuring in the content of that attitude. What we need, then, is a suitably non-normative account of this attitude. However, providing a non-normative account of attitudes with normative content is precisely what expressivist theories are for. The solution to the worry is therefore simply to apply our preferred expressivist theory to the attitude in question.

Now, it might seem that expressivists face a problem here. This is because, for any proposition  $p$  with normative content, expressivist theories tell us only what it is to *believe* or *judge* that  $p$ . They do not tell us what it is to *desire*, *hope*, *doubt*, *fear*, and so on, that  $p$ . Hence, expressivists face the 'many attitudes problem,' since they apparently lack any account of normative attitudes besides belief.<sup>17</sup> With this in mind, we can see that the normativity worry described above boils down to a particular instance of the many attitudes problem. For what we are after is an expressivist account of a particular kind of normative desire—desiring that which it is rational to desire in some context or other.

The solution to the normativity worry therefore lies in the solution to the many attitudes problem more generally. And fortunately, a solution is ready to hand. As [Beddor \(2020\)](#) and [Baker \(2021\)](#) both observe, given plausible assumptions within the philosophy of mind, expressivists can explain different kinds of attitudes in familiar ways. For instance, given a functionalist view of propositional attitudes, desiring  $p$  is a dispositional state individuated by its causal-functional connections to certain inputs and outputs. Simplifying somewhat, we might suppose that *desiring*  $p$  is that state which produces certain actions or intentions (its outputs) when combined with beliefs about what is likely to bring about  $p$  (its inputs). In this way, the content of the desire is fixed in terms of its causal-functional connections to beliefs involving the same content. Thus, if  $p$  has normative content, desiring  $p$  is explained in terms of its causal-functional connections to beliefs about  $p$ . Crucially, since

<sup>17</sup> The name comes from [Schroeder \(2010\)](#) but was first raised by [Rosen \(1998\)](#). Why isn't this also a problem for representationalist theories? Because representationalist theories can explain other normative attitudes in terms of bearing different relations to the same proposition. Since the proposition in question has normative content, this strategy is not open to expressivists, who eschew an explanatory role for normative content.

expressivists already have a suitably non-normative account of what it is to believe  $p$ , our account of desiring  $p$  can be given in non-normative terms. Specifically, to desire  $p$  is to stand in certain casual-functional connections to one's beliefs about  $p$ , which are explained in non-normative terms.

Thus, desiring that which it is rational to desire is to be in a state that is causally connected in the right way to one's beliefs about the rationality of desires.<sup>18</sup> So far, so normative. But once we explain those beliefs in terms of the norms, standards, or values one accepts, then we have provided a suitably non-normative account of what it is to desire that which it is rational to desire. Thus, for example, if I desire that things go well for you, we might explain this in terms of its characteristic inputs—the norms I accept about what to want for you insofar as I care for you, say—and its characteristic outputs—my wanting these things for you.<sup>19</sup> And *mutatis mutandis* for other propositional attitudes with the same content.

Let's now turn to the new circularity worry. The worry arises because we have explained the nature of care by appealing to an attitude which is itself explained in terms of care. Specifically, care is explained in terms of desiring that which it is rational to desire insofar as one *cares* for  $x$ . Moreover, if many of the components of care involve prudential notions, as in Jaworska's characterization, and if the prudential circularity for each of these components is removed by applying the rational care analysis in the manner described above, then it follows that all such components of care will be partly explained in terms of care. So it might seem that our account is viciously circular after all, perhaps rampantly so, even if not for the reasons originally thought.

As before, the solution to the problem lies in first observing that 'care' appears in the content of an attitude and then specifying the role it plays in individuating the attitude. To begin, given that 'care' denotes a cluster of attitudes, we can remove any reference to 'care' and instead appeal to the set of attitudes that constitute the cluster. Thus, to desire  $p$  insofar as one cares for  $x$  is to desire  $p$  insofar as one possesses attitudes  $A_1, \dots, A_n$ , (or a sufficient subset thereof) where these are the joys, frustrations, fears, hopes, and so on, that together constitute caring for

<sup>18</sup> As Baker (2021) argues, expressivists need not be committed to any particular account of the relevant functional relations. Though see Beddor (2020) for specific suggestions.

<sup>19</sup> There is a question here for expressivists concerning how exactly normative judgments 'motivate' or govern our attitudes (for example, desire) and not only our actions. Perhaps the relevant outputs will not include the attitude itself, but rather actions or feelings that somehow serve to bring the attitude about. However, this is a question for any internalist theory that allows for reasons for attitudes rather than specific to the present account.



$x$ . While this removes any reference to care in our characterization of this attitude, it still exhibits circularity. Firstly, this is because among the attitudes mentioned in our characterization is the very desire we are trying to explain. Secondly, insofar as this desire is a component of care, it will also be mentioned in our characterization of every other attitude in  $A_1, \dots, A_n$  that has prudential content. However, rather than attempt to remove this circularity, the correct response is to note that it is non-viciously circular and thus unproblematic. For there is nothing viciously circular in having a desire for something *insofar as this desire forms part of a complex cluster of interrelated hopes, fears, joys, and so on*.

To care for  $x$ , then, is to possess a complex cluster of causally related desires, joys, fears, and so on, towards  $x$ . To desire  $p$  insofar as one cares for  $x$  is to desire  $p$  on the condition that one instantiates this web of attitudes. A question one might have about the proposed account is how demanding the possession conditions on caring are. Must one instantiate the whole cluster or only part of it? If only part of it, how much? It seems plausible that we can care for something even if we lack some of the relevant caring attitudes. For supposing we possessed all but one of the relevant attitudes towards some object, it seems that attributing care for that object would still be appropriate. A less clear case might be one in which we realize we care for something only when we lose that thing. This might be because some of the relevant dispositions have been masked, but it might simply be that while we possess the relevant attitudes related to that thing's loss, we lack those related to its flourishing. Rather than pose a difficulty for our view, however, what this brings out is that caring is an attitude that comes in degrees. How much one cares for something will often be explained by the strength of the relevant desires, fears, joys, and so on, towards the object of care. But it might also be explained by how much of the cluster one instantiates. Precisely where the threshold lies for care attributions will no doubt be vague, but this is a feature of many kinds of attitudes.

Combining the responses to the two worries, we now have the following account of what it is to desire that something fares well:

$S$  desires that  $x$  fares well if and only if  $S$ 's desire stands in the right functional relations to  $S$ 's beliefs about what it is rational to want insofar as one possesses attitudes  $A_1, \dots, A_n$  towards  $x$ .

Thus, for instance, if desires are partly defined by the way they combine with beliefs to produce actions, then desiring that  $x$  does well will involve a disposition to bring about what one believes it is rational to want insofar as one possesses this desire and other caring attitudes

towards  $x$ . If we then give an expressivist account of this belief, we provide a suitably non-normative, non-circular account of one component of care. The other components are explained in the same way but in terms of whatever functional relations to  $S$ 's beliefs (and other attitudes) are constitutive of those attitudes.<sup>20</sup>

3.2.4 *The point of prudence* Imagine, then, a disagreement between a prudential hedonist and a prudential ascetic concerning an important life decision of a mutual friend. Whereas the hedonist thinks it would be best for the friend to choose the path of most pleasure, the ascetic thinks it would be best for the friend to choose the path that renounces pleasure. What is this disagreement about? According to the expressivist rational care theory, the disagreement concerns what norms or values to accept in relation to what to want insofar as they care for their friend, which is to say insofar as they in fact have this desire and other related attitudes governed by the very same norms or values. The hedonist accepts norms or values that prescribe wanting pleasure for those we care about, and insofar as the hedonist cares for the friend, they will desire pleasure for the friend. The ascetic accepts norms or values that proscribe pleasure for those we care about, and insofar as the ascetic cares for the friend, they will desire that the friend foregoes pleasure. Thus, while each person wants something different for their friend, both cases are manifestations of care. While they share the *de dicto* desire for their friend's good, their differing desires are a consequence of their differing beliefs about what their friend's good consists in. In this way, caring for a person involves being guided by one's beliefs about how to rationally care for that person.

Many details remain to be filled in and there are many complexities not considered here that a mature theory would need to accommodate. The point, however, has not been to develop a fully worked-out account, but to show how it is possible to provide a non-circular, non-normative characterization of what it is to care, even if caring essentially involves responding to another's well-being. I conclude this section of the paper by offering an independently motivated, speculative story about the role of care that is of a piece with the kind of explanations expressivists give about the practical role of normative attitudes more generally.

<sup>20</sup> See [Beddor \(2020\)](#) and [Baker \(2021\)](#) for further discussion of some of the details. While I have focused on how we can explain this desire within a functionalist framework, the argument should generalize to any approach that explains propositional attitudes in terms of their constitutive relations to other attitudes. For instance, these might be relations of constitutive rationality rather than causal-functional relations; compare the interpretationist approach proposed in [Brown \(2022\)](#) and the discussion of the many attitudes problem therein.

From as early as the second year, human beings already begin to exhibit caring behaviours in response to the perceived distress of others (Zahn-Waxler et al. 1992). This arguably occurs via the activation of innate empathy-arousing mechanisms that cause us to have feelings more congruent with the other's situation than our own (Hoffman 2000). If the development of our caring responses ended there, then there would be no grounds for thinking that care is a normative attitude. Rather, care would simply be a response to the perceived distress of another. However, as we develop in late childhood and adolescence, we begin to respond empathically not simply to another's distress, but to their 'life situation' more generally (Hoffman 2000, p. 80). It is at this point of development that it begins to make sense to ask which features of another's life situation are appropriate objects of one's empathic responses. For there is no descriptively given criterion for which features of their life situation are the correct object of one's responses. But to ask this is to ask a normative question. And it is precisely our prudential concepts that allow us to pose such questions. This is what our prudential concepts are *for*.

Thus, if it is part of our mature attitude of care to be able to reflect on and revise our implicit or explicit views about the features of another's life situation it is appropriate to respond to from the perspective of care, then it would seem that care is itself a normative attitude. However, as long as the normative concept of 'appropriateness' involved in this attitude can be given an expressivist interpretation, there is nothing here incompatible with expressivism. Admittedly, this story is somewhat speculative. However, it is empirically informed and plausible enough given general facts about human development and interaction. This shows that the expressivist rational care theory can be offered not merely as an *ad hoc* response to the challenge of prudential normativity. It is independently motivated by exactly the kind of story that expressivists want to tell about normative thought and discussion more generally: namely, that normative thought and discussion is fundamentally explained in terms of its motivational, cooperative, and coordinative role in allowing human beings to live together.

#### 4. Conclusion

If prudential thought and discussion is normative, then any mature meta-normative theory must explain what makes it normative and what distinguishes it from other kinds of normative inquiry. Our theory answered the second question in terms of its being distinctively about the rationality of caring attitudes. It answered the first question in terms of whatever

expressivist explanation we give of the attitude expressed by ‘rational’. While it would be nice if the fitting attitude strategy generalized to other varieties of normativity, whether it does so remains to be seen. However, I hope that the above discussion merits further investigation in this direction.

While I have defended the rational care view from what I take to be its most pressing objection, there will of course be other objections to the approach that I have not considered here. Moreover, I have not considered how the approach fares with respect to competing expressivist approaches. This is partly because of the current lack of expressivist approaches, but it is mostly in order to focus on the constructive task of providing an expressivist account of prudential thought and discussion. While our account ultimately needs to be assessed against its competitors, I leave this task for elsewhere.

That said, let me briefly comment on [Beddor’s \(2021\)](#) proposal that expressivists should individuate normative domains in terms of distinctive kinds of approval and disapproval. In the current context, the idea is that while all normative claims express (dis)approval of some kind, prudential claims express specifically prudential (dis)approval. Beddor then offers a number of ways in which one might explain prudential (dis)approval. However, every strategy Beddor offers involves appealing to other attitudes with prudential content. For example, he suggests that prudential disapproval might be disapproval based in ‘the desire for the wellbeing of a particular agent’ (2021, p. 29). He also suggests that prudential disapproval is functionally connected to prudential blame, which is blame based on prudential considerations (2021, pp. 29–30). In both cases, Beddor fails to offer a non-normative, non-circular characterization of the relevant attitudes. So the proposal, at least as it is developed there, fails to offer a genuine alternative for expressivists.

By contrast, the expressivist rational care theory provides a non-normative, non-circular characterization of prudential thought and discussion. In doing so, it solves the challenge of prudential normativity for expressivism and the circularity problem for the rational care theory. Thus, while the marriage may be one of convenience, I am hopeful it is a happy one.<sup>21</sup>

<sup>21</sup> I would like to thank the following people for many helpful comments on many previous versions of this paper: Matthew Chrisman, Guy Fletcher, Sebastian Köhler, James Laing, James Lewis, Sophie Potter, Michael Ridge, Christine Tiefensee, and the reviewers and editors at *MIND*. Special thanks to Guy Fletcher for first suggesting the topic to me. The paper also benefited from feedback from audiences at the Frankfurt School of Finance & Management, Sapienza Università di Roma, Universitat de Barcelona, Universität Bonn, and University of Cyprus. It was supported by funding from the British Academy (Grant No. PF21\210089) and the Alexander von Humboldt Foundation.

## References

- Anderson, Elizabeth (1993). *Value in Ethics and Economics*. Harvard University Press.
- Baker, Derek (2021). Deflating the Many Attitudes Problem. *Philosophical Quarterly* 72 (1): 1-18.
- Beddor, Bob (2020). A Solution to the Many Attitudes Problem. *Philosophical Studies* 177 (9): 2789-2813.
- Beddor, Bob (2021). Moral and Epistemic Evaluations: A Unified Treatment. *Philosophical Perspectives* 35 (1): 23-49.
- Blackburn, Simon (1984). *Spreading the Word: Groundings in the Philosophy of Language*. Clarendon Press.
- Blackburn, Simon (1998). *Ruling Passions: A Theory of Practical Reasoning*. Oxford University Press UK.
- Brown, James L. D. (2022). Interpretative Expressivism: A Theory of Normative Belief. *Philosophical Studies* 179: 1-20.
- Brown, James L. D. (2024). On Scepticism About Ought *Simpliciter*. *Australasian Journal of Philosophy* 102 (2): 479-511.
- Campbell, Stephen M. (2016). The Concept of Well-Being. In Fletcher, Guy (ed.) *The Routledge Handbook of Philosophy of Well-Being*. Routledge.
- Darwall, Stephen (2002). *Welfare and Rational Care*. Princeton University Press.
- Darwall, Stephen (2006). Reply to Feldman, Hurka, and Rosati. *Philosophical Studies* 130 (3): 637-658.
- Feldman, Fred (2006). Welfare and Rational Care. *Philosophical Studies* 130 (3): 585-601.
- Fletcher, Guy (2019). Taking Prudence Seriously. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics* 14. Oxford University Press.
- Fletcher, Guy (2021). *Dear Prudence*. Oxford University Press.
- Gibbard, Allan (1990). *Wise Choices, Apt Feelings*. Harvard University Press.
- Gibbard, Allan (2003). *Thinking How to Live*. Harvard University Press.
- Gibbard, Allan (2006). Moral Feelings and Moral Concepts. In Russ Shafer-Landau (ed.) *Oxford Studies in Metaethics 1*. Oxford University Press.
- Hills, Alison (forthcoming). Aesthetic Obligation. In Simon Kirchin (ed.) *The Future of Normativity*. Oxford University Press.
- Hoffman, Martin L. (2000). *Empathy and Moral Development*. Cambridge University Press.
- Jaworska, Agnieszka (2007). Caring and internality. *Philosophy and Phenomenological Research* 74 (3): 529-568.

- Kagan, Shelly (1994). Me and My Life. *Proceedings of the Aristotelian Society* 94: 309-324.
- Köhler, Sebastian (2013). Do Expressivists Have an Attitude Problem? *Ethics* 123 (3): 479-507.
- Köhler, Sebastian (2021). Normative Disagreement: A Functional Account for Inferentialists. *Philosophical Studies* 178 (2): 617-637.
- Kratzer, Angelika (1977). What 'Must' and 'Can' Must and Can Mean. *Linguistics and Philosophy* 1: 337-355
- Lenman, James (2024). *The Possibility of Moral Community*. Oxford University Press.
- Lin, Eden (2022). Well-Being, Part 2: Theories of Well-Being. *Philosophy Compass* 17 (2): 1-15.
- Maguire, Barry & Woods, Jack (2020). The Game of Belief. *Philosophical Review* 129 (2): 211-249.
- Maynard, Jonathan Leader & Worsnip, Alex (2018). Is There a Distinctively Political Normativity? *Ethics* 128 (4): 756-787.
- Miller, Alexander (2013). *Contemporary Metaethics: An Introduction*. Polity.
- Plunkett, D., Shapiro, Scott & Toh, Kevin (eds.) (2019). *Dimensions of Normativity: New Essays on Metaethics and Jurisprudence*. Oxford: Oxford University Press.
- Rabinowicz, Wlodek (2013). Value, Fitting-Attitude Account of. In Hugh LaFollette (ed.), *The International Encyclopedia of Ethics*. Wiley-Blackwell.
- Railton, Peter (1989). Naturalism and Prescriptivity. *Social Philosophy and Policy* 7 (1): 151-174.
- Ridge, Michael (2014). *Impassioned Belief*. Oxford University Press.
- Ridge, Michael (2024). Normativity, Prudence and Welfare. *Philosophical Studies* 181 (5): 1213-1235.
- Rønnow-Rasmussen, Toni (2011). *Personal Value*. Oxford University Press.
- Rowland, Richard (2016). Reasons as the Unity Among the Varieties of Goodness. *Pacific Philosophical Quarterly* 97 (2): 200-227.
- Schroeder, Mark Andrew (2008). *Being For*. Oxford University Press.
- Schroeder, Mark (2010). *Noncognitivism in Ethics*. Routledge.
- Shah, Nishi (2004). Welfare and Rational Care. *Philosophical Review* 113 (4): 577-582.
- Sinclair, Neil (2021). *Practical Expressivism*. Oxford University Press.
- Smith, Michael (2001). Some Not-Much-Discussed Problems for Non-Cognitivism in Ethics. *Ratio* 14 (2): 93-115.

- Smith, Michael (2002). Which Passions Rule? *Philosophy and Phenomenological Research* 65 (1): 157-163.
- Wedgwood, Ralph (2007). *The Nature of Normativity*. Oxford University Press.
- Wodak, Daniel (2017). Expressivism and Varieties of Normativity. In Russ Shafer-Landau (ed.) *Oxford Studies in Metaethics* 12. Oxford University Press.
- Worsnip, Alex (2021). *Fitting Things Together: Coherence and the Demands of Structural Rationality*. New York: Oxford University Press.
- Zahn-Waxler, C, Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of Concern for Others. *Developmental Psychology* 28: 126-136.