

## 5

# Internalism and Prudential Value

C5

*Jennifer Hawkins*

C5.P1 Many philosophers accept that prudential value is a distinct type of value, different both from moral value and also from general non-moral goodness.<sup>1</sup> Those who (like myself) accept the category of prudential value think of it as the type of value at stake when we say that something is non-morally good *for* a particular person. It is the kind of value theories of welfare or well-being are about. It is what we seek for ourselves on those occasions when we seek our own self-interest. And it is the kind of value we seek to realize for others when we try to benefit them. There are many interesting and (relatively) underexplored foundational questions about prudential value. This paper is concerned with whether or not some form of existence internalism (defined below) holds for prudential value, and if so what it amounts to.<sup>2</sup>

C5.P2 Existence internalism is the idea that there is a necessary connection between something's having a certain positive normative status (its being right or its being intrinsically prudentially good) and its ability to elicit positive psychological responses from human subjects.<sup>3</sup> Existence internalism is typically contrasted with judgment internalism, which is a claim about what follows, as a matter of conceptual content, from sincere

<sup>1</sup> Philosophers differ in how they carve up value. For example, Moore (1903) and Scanlon (1998) are skeptical of the very idea of prudential value that is relative to individuals. However, they are not skeptical of the idea of general, non-relational non-moral goodness. Others who discuss general, non-relational goodness take the notion of intrinsic goodness to be primarily a *moral* notion (Zimmerman 2015).

<sup>2</sup> The idea that some form of internalism must hold for prudential welfare is commonly expressed, but there are relatively few explicit discussions of what such a constraint might amount to. At the time this essay was written in 2014, the main discussions of prudential internalism were Rosati (1996), Velleman (2000), and Sarch (2011). Since then there have been several new additions to the literature. See, in particular, Dorsey (2017) and Heathwood (ms.).

<sup>3</sup> The term 'existence internalism' and its close relative 'judgment internalism' come from Darwall (1983: 54). A similar distinction with different labels occurs in Brink (1989: 40).

judgments about value.<sup>4</sup> Existence internalism, however, is a metaphysical thesis, not a conceptual one.

C5.P3 Put another way, it is the claim that facts about human psychological responsiveness constrain the metaphysics of value in particular ways. Attempts to specify existence internalism for a particular domain of value are attempts to say what the metaphysical constraints are for that type of value. Peter Railton characterizes the task in terms of avoiding “alienness.” He writes, “What is intrinsically valuable for a person must have a connection with what he would find to some degree compelling or attractive, at least if he were rational or aware. It would be an intolerably alienated conception of someone’s good to imagine that it might fail in any such way to engage him.”<sup>5</sup> However, there remain different views about what would suffice to avoid alienness. My aim is to introduce and defend as viable a new account of what such constraints might look like.

C5.P4 Moral existence internalism is typically formulated as a claim about *motivation*. This reveals an important sense in which discussions of prudential value differ significantly from discussions of moral value. In the moral case, we are typically invited to imagine an agent who has the opportunity to act rightly or wrongly. The question is whether the moral facts are such that her mere awareness of them would give her (at least some, perhaps defeasible) motivation to do the right thing. But this scenario is *prospective*. Talk of motivation makes sense only when action is required, and action is only *sometimes* required in relation to prudential value. Sometimes we are in *pursuit* of the good. But at other times we talk about goodness that is realized, and here talk of motivation seems out of place. When we talk about goods that a person ‘has’ or ‘possesses’ we need different language. Rather than speak of motivation we speak of *caring about* the good or *appreciating* the good or some other *contemporaneous* positive response.

C5.P5 Several theorists have recognized that when we theorize about prudential value we need a broader psychological notion than just *motivation*.<sup>6</sup> However, this issue of terminology points to something with more than merely terminological significance. It points to a modal distinction between types of prudential facts—a distinction between goodness that is *currently realized* and goodness that is *merely possible*; although no one has made much of this modal distinction, I will do so here. In what follows, I defend the idea that

<sup>4</sup> Smith (1994) defends *moral* judgment internalism. For detailed surveys of the many forms metaethical internalism can take and the relevance for particular debates, see Van Roojen (2015) and Rosati (2016).

<sup>5</sup> Railton (1986a).

<sup>6</sup> In addition to Rosati (1996), see Velleman (2000), Dorsey (2017), and Heathwood (ms.).

prudential existence internalism should be understood simply as a constraint on facts about realized goodness. In those cases where talk of motivation might make sense—cases involving possible good—I argue that we should be motivational externalists.

C5.P6 In Section 5.1, I first briefly describe what I call “the structure of good.” The aim is simply to make explicit the important relationships between different kinds of goodness claims. In Section 5.2, I then introduce and defend the coherence of my particular version of internalism. I don’t pretend to offer a direct argument for the plausibility of prudential existence internalism itself.<sup>7</sup> Instead, since many theorists of well-being accept internalism, I aim to demonstrate the coherence and plausibility of a *particular way of formulating* existence internalism for prudential value. In Section 5.3, I examine the most influential alternative formulation of the internalism constraint, Rosati’s two-tier internalism, and argue that my account is preferable to hers.

C5.S1

## 5.1. THE STRUCTURE OF GOOD

C5.P7 Ordinary talk about what is or would be good for a person is usually talk about *prudential goods*. We seek to realize various kinds of intrinsic prudential goods in our lives. I adopt an understanding of intrinsic value as final value, such that its natural opposite is instrumental value.<sup>8</sup> However, it is worth noting that the word ‘intrinsic,’ when used to signal final value, could also apply to the kind of value had by a good life.<sup>9</sup> Here, I only use the word for prudentially valuable items within a life. Furthermore, I take no stance here on what kinds of things may count as basic intrinsic prudential goods.

<sup>7</sup> Some theorists do argue for it. See Rosati (1996), Velleman (2000), and Dorsey (2017). Sarch (2011) offers important criticisms of the arguments in Velleman and Rosati. Dorsey (2017) argues that we should not accept prudential existence internalism simply because we find it highly intuitive. There are, he claims, far too many intuitions that pull in the opposite direction as well, such that we must have some substantive reason for holding onto internalism. I think he overstates his point, however. Those who find internalism compelling rarely find the alleged counterexamples to internalism compelling as well. Rather this is one of the places in philosophy where (sadly) intuitions seem to fall on different sides of a seemingly unbridgeable divide.

<sup>8</sup> See Korsgaard (1983) and Kagan (1998). Intrinsic value can refer either to the value something has based only on its intrinsic properties (the contrast is extrinsic value), or it can refer to the kind of value something has when we value it for itself, as a final end (the contrast being instrumental value). Kagan defends the coherence of retaining the label “intrinsic value” for the concept of final value, which is the approach I adopt. For more about the concept of intrinsic value, see Zimmerman (2015).

<sup>9</sup> This point is made by Feldman (2000).

Everything I say can be adapted (sometimes with a little tweaking of my examples) to the full range of views about prudential goods.

C5.P8 Basic, intrinsic, prudential goods are good for us at a specific time or times.<sup>10</sup> If riding a bicycle is intrinsically good for me, then, plausibly, it is good for me *while I am riding*. Some goods are temporally brief; others last most of a lifetime. It is important to mark a distinction between goods that are currently realized for a person (currently adding value to her life) and goods that could be realized (depending on how things go). We must constantly be aware of the modal distinction between realized good and possible good.

C5.P9 Because I maintain that prudential goods are the building blocks of life value, my view is unabashedly additive. What makes a life good are the goods in it. Of course, some theorists reject additive views.<sup>11</sup> However, I suspect that many people are suspicious of additive accounts because they focus on rather simple versions of additivism.

C5.P10 For additivists like myself, talk of a good life depends on the notion of overall value, or net value. Importantly, a life can be said to be good in two distinct ways.<sup>12</sup> In one sense, a life is good if it has net positive value, if the intrinsic goods in that life outweigh the intrinsic bads, even if only by a little. Something like this seems to be what (some) philosophers have in mind when they talk about “a life worth living.” However, a different sense of good captures the sense in which we all *aspire* to have a good life. A good life in this sense is one that satisfies a high threshold. It is a life that has *a fair amount of net positive value*. It is something to aim at. Lives like this are “better than most.”

C5.P11 Insofar as we exercise agency in pursuit of our good, what concerns us most are *good choices*. To understand choices, we must consider the structure of possible future good. For any given individual there are many possible continuations of her life. What counts as the *best* continuation depends on what the set of possibilities are. The *best* continuation is the one with the greatest net value. However, it is important to remember that the best continuation of a life

<sup>10</sup> Not everyone will agree that goods must be realized at determinate times (e.g. Nagel 1970). I argue for the idea that prudential goods must be realized at particular times in Hawkins (2014).

<sup>11</sup> A certain sub-set of theorists reject additive views because they take seriously the idea that the shape of a life matters prudentially. They tend to assume that this rules out additivity (or as some say “intra-life aggregation”). See e.g. Slote (1984), Velleman (1991), Temkin (2012), Glasgow (2013). Dorsey (2015) demonstrates that, depending on *why* you think shape matters, taking shape seriously may be compatible with aggregation. I am more willing than many to simply downplay the importance of shape of life considerations. On this see Hawkins (2014).

<sup>12</sup> I find it intuitive to distinguish (as I do in the text) between a minimal sense of good and a more robust, aspirational sense that depends on a high threshold. However, sometimes the two are collapsed. For example, Sidgwick (1907) equates welfare with happiness and defines happiness as a “surplus of pleasure over pain.”

(which means ‘best’ relative to the options) may not be good in either of the senses sketched above: it may not have *positive* net value, and it may not be good in the aspirational sense of a life “better than most.”

C5.P12 The *best* choice for an individual at a time is the one that moves her in the direction of realizing the best continuation of her life. A *good* choice is one that moves her towards one of the better continuations in the set of possible continuations. In ordinary speech, evaluations of choices and outcomes often get blurred. For example, we are often reluctant to call a choice good if the outcome it produces is not (even though it leads to the least worst outcome). And sometimes we call a choice good if it helps us to realize something good *immediately*, even though the further consequences of the choice are not good. In what follows, however, I stick to the use of “good choice” defined here.

C5.P13 Sometimes realizing a good continuation of a life requires the subject to make dramatic changes—changes of direction or goal. It may require new skills or training. It may even require changing aspects of herself. This is important, because goodness facts that lie on the other side of great change are the hardest for us to “see.” It could thus be true of an individual that the best choice for her is one she does not recognize as good at all. Like intrinsic prudential goods, choices have value at a time. But their value is instrumental, depending on how the choice relates to the possibilities the subject confronts.

C5.S2

## 5.2. EXISTENCE INTERNALISM AND PRUDENTIAL VALUE

C5.P14 Let us now consider potential internalist constraints on goodness facts. Recall that existence internalism is the claim that there is a necessary connection between, on the one hand, normative facts (in this case *prudential* facts) and on the other hand, human psychological responses. What might this connection be?

C5.P15 In the case of prudential value there is a rather simple suggestion one can make. Unsurprisingly, being simple it has problems. However, these are more easily dealt with than typically supposed. I shall consider and respond to these problems along the way, some in this section and some later. But for the moment, reflection on the simple suggestion is a good way to start.

C5.P16 What I call the simple suggestion is best brought out through examples.<sup>13</sup> Imagine a woman named Samia who is a newspaper writer. Various people

<sup>13</sup> The following example presupposes a non-mental state theory of welfare since the activity of writing is being considered as a candidate for *intrinsic goodness*. Nothing much

claim that being a writer is intrinsically good *for* Samia. They mean that it is good for her now, not that it will be good for her at some point in the future, nor that it would be good for her if only she were different in some way. Because they mean it is good for her now—as she is—they are making a claim about *realized intrinsic prudential goodness*.

C5.P17 As we learn more about her case, the facts cast doubt on their claim. First, we discover that writing and the writing life do not give Samia *any* pleasure.<sup>14</sup> Next, we consider whether it makes her happy. After all, people can sometimes be made happy by activities or projects that they would not exactly describe as *pleasurable*. But the evidence suggests that Samia is not happy. Writing leaves her feeling depressed and drained. Sometimes people pursue things they *value* even though the pursuit doesn't fit their affective profile very well. In such cases, despite how the pursuit makes them feel, they think *the thing they are pursuing* has value. It is because they value *it* that they persevere. But Samia doesn't value writing either. She doesn't think it is important in its own right. In fact, she doesn't see any value in it at all. For whatever reason, it is just the job she has right now and she has to do it to make a living.

C5.P18 At this point it becomes incredibly hard to believe that writing is intrinsically good for Samia. This suggests a general type of necessary condition on *realized intrinsic goodness*. Call this the SIMPLE CONSTRAINT (SC):

C5.P19 SC: A necessary condition on something X being a basic intrinsic prudential good for a subject, S, at a time T, is that S must have some kind of positive response to X at T.

C5.P20 Here, I intend positive response to be construed quite broadly to include positive sensations, feelings, affective states, and positive evaluative attitudes.<sup>15</sup> However, positive responses must be psychological and at least partly conscious. It is also important to note that the condition is framed as necessary, but it says nothing about the sufficiency conditions for good.

C5.P21 One problem with SC emerges quickly. It has to do with our understanding of 'response' and whether it is flexible enough to cover cases of temporally extended value and the attitudes that ground such value. Suppose (unlike a hedonist) I think that relationships can have intrinsic prudential value. A relationship might last for a long time and have value for

turns on this as I hope will be clear by the end. But a concrete example has to make some sort of assumption about welfare.

<sup>14</sup> Even non-hedonists may think that pleasure is a sign of the presence of value.

<sup>15</sup> My intention is to be as broad as possible within the range of positively valenced psychological reactions. Heathwood (ms.) similarly argues for the need to define the psychological side of an internalist constraint broadly. We thus differ from Dorsey (2017) who defines the constraint in terms of valuing attitudes.

me during all of that time. But presumably I am not having a positive response to this relationship (or the person I am related to) at each of the times when (intuitively) it is good for me. What seems likely is that the relationship becomes good when I first develop a certain attitude of valuing towards it. But this attitude will not be at the forefront of awareness from then on.

C5.P22 Of course there is a sense in which, all through that time, I ‘have’ the valuing attitude. This is a dispositional sense. Most beliefs and evaluative attitudes require a dispositional analysis. For years I have had a belief about where I was born. If you ask me if I was born there I will say yes. But I hardly ever think about this fact, despite believing it. Only rarely does this content rise to the level of occurrent thought.

C5.P23 Evaluative attitudes are more complex than ordinary non-evaluative beliefs. It might call into doubt my valuing of X if I rarely thought about X. If I value X, occurrent positive thoughts about X must occur, at least periodically. Moreover, if X is something (or someone) I interact with regularly, one would expect me to have various kinds of positive response to X on a regular basis. It was (at least in part) the absence of such regular, positive responses to writing that raised doubts about Samia. It is, of course, entirely unclear how many responses are required or how often they must occur. But the details need not concern us here as long as we can agree that it is too strong to say that positive response must be constant, and too weak to require that positive response only occur once. I suggest we talk about valuing and other ongoing positive attitudes as ‘positive engagements,’ in recognition of the fact that such attitudes, though dispositional, will with some regularity become manifest in occurrent positive responses. Then we can make the following amendment:

C5.P24 SC2: A necessary requirement on X’s being a basic, intrinsic prudential good for S at T is that S must either have a positive response to X at T, or S must have an ongoing positive engagement with X that still holds at T.

C5.P25 However, a second problem follows quickly in the wake of the first. For SC2 seems to require that all good things *enter the awareness of a subject*. Most contemporary theorists of welfare reject what James Griffin labeled “the experience requirement,” the idea that prudential goodness (or badness) must enter our experience or our awareness.<sup>16</sup> They insist that intrinsic prudential value can be realized even when the subject is ignorant of this. The most straightforward example of theories that allow this are traditional versions of desire fulfillment theory that allow that the satisfaction of a desire

<sup>16</sup> Griffin (1986).

(an event in the world) is good for a person even if she doesn't ever know it has been satisfied.

C5.P26 This is a problem; however, it is important to keep certain goals in mind. The aim is to articulate a very general internalism requirement that focuses on *realized intrinsic prudential good*. Any candidate must minimally explain our intuitions about cases like Samia. But we can accomplish these dual goals without at the same time building in an experience requirement. We just change SC2 to SC3:

C5.P27 SC3: A necessary condition on something, X, being a basic, intrinsic prudential good for a subject, S, at a time T, is that:

C5.P28 If S is aware of X at T, S must either have a positive response to X at T, or S must have an ongoing positive engagement with X that still holds at T, *or*

C5.P29 If S is unaware of X at T, it must be true that S would have had a positive response to X at T if she had been aware of X at T.

C5.P30 This formulation leaves room for the possibility that something, X, might be good for S at T, even though S does not respond positively to it then *because S knows nothing about it*. It holds onto the idea (which seems key) that a certain kind of psychological relation must obtain between S and X at the very time X is good for S. What is relinquished is the idea that this relation must involve a direct experience of the good thing. The right kind of psychological relations can obtain at a time as long as certain psychological dispositions for response obtain at that time.<sup>17</sup> It is also important that the counterfactual element here is minimal. No appeal is made to how S would respond to X in conditions extremely different from the ones she is currently in. The counterfactual appeals to a situation exactly like the real one but for one difference: S's awareness of X. Finally, such an account can still easily handle the more ordinary cases in which someone *is* aware of a putative good. Writing is not good for Samia, since Samia clearly is aware of writing yet has no positive responses to it and no ongoing positive engagement with it.

C5.P31 Some theorists may still find SC3 problematic. Even though there is no requirement in SC3 that S actually respond positively to X, there is a requirement that S have certain *response dispositions* toward X at T. Some theorists think that once you give up the experience requirement there is no

<sup>17</sup> The current framing of the constraint assumes that the intrinsically good thing must be the *object* of a psychological response. As will become clear in later discussion, I also allow that a theorist could modify SC slightly to include cases in which the psychological response is itself the intrinsic good. Heathwood (ms.) also discusses this.



reason to place restrictions on the subject's psychology at the time when goodness is realized.<sup>18</sup> But this is mistaken. Without some such requirement it is far too easy to generate bizarre cases in which a person is benefited at a time by something she would (if she actually knew about it) view as deeply bad. The constraint we are articulating requires (as a necessary condition only) a kind of fit between current psychological outlook and goodness facts. No doubt different theorists will be tempted to explain in different ways why such fit is important. But whatever explanation one gives, it is hard to see why it will not apply just as well in the case where the subject is ignorant of the putative good.

c5.P32 In what follows, I shall continue to use the phrase “simple constraint” but I will now take it to refer to SC3.

### c5.S3 5.2.1. Internalism and Theories of Welfare

c5.P33 In addition to whatever surface plausibility the simple constraint has, it gains some indirect support from the fact that many theorists of welfare, particularly those of a subjectivist bent, seem to be—at least implicitly—concerned with preserving something like it.

#### c5.S4 5.2.1.1. *Hedonism*

c5.P34 Consider first classical hedonism according to which pleasure is the only basic intrinsic good and pain the only basic intrinsic bad. One natural way of understanding the allure of hedonism is to note that pleasure is a mental state that is, when experienced, always *welcomed*, at least to some degree, by the person whose pleasure it is. It is, of course, possible to have mixed reactions to pleasure. I might feel pleasure, but because I think my feeling pleasure is inappropriate, I might also feel shame, or embarrassment. In short, I can both feel pleasure at a time (and when feeling it welcome it) and yet at a different, higher level of my psychology, adopt a disapproving attitude towards my own pleasure and so desire not to be experiencing it.

c5.P35 I suggest we think of this ‘welcoming’ response as an involuntary affective response, one that operates at a different level from more cognitively sophisticated attitudes. It is this immediate, low-level reaction to pleasure—what I am calling the welcoming response—that seems to underwrite our sense that experiences of pleasure satisfy SC3. Moreover, the welcoming response is something we reliably have whenever we experience pleasure, and it lasts only for as long as the pleasure lasts. Thus pleasure

<sup>18</sup> Derek Parfit (1984: 495) seems to suggest this.

naturally and easily satisfies the simple constraint: for the hedonist, pleasure is good whenever S experiences it, and whenever S experiences it, S has a positive (welcoming) response to it at that time.

C5.P36 Someone might object that whether or not pleasure satisfies the simple constraint depends on the nature of pleasure.<sup>19</sup> There is something to this. To see why, consider two views—distinguished by L. W. Sumner—about the nature of pleasure.<sup>20</sup> On internalist views, pleasure is simply a particular type of sensation. What various pleasures have in common is “their positive feeling tone: an intrinsic, unanalyzable quality of pleasantness that is present to a greater or lesser degree in all of them.”<sup>21</sup> Externalist views, on the other hand, define pleasure in terms of some attitude or response that an agent has towards a sensation. The various sensations might be quite diverse, but what makes them pleasures is the subject’s attitude towards them.<sup>22</sup> I find it helpful to think of such theories as bundle theories, because on such views pleasure is a bundle with two elements: a sensation *and* a certain kind of positive response to that sensation.

C5.P37 Both internalists and externalists typically acknowledge the existence of positive response in cases of pleasure. They differ merely in whether they view that response as a reaction *to* pleasure, or as a *part of* pleasure.

C5.P38 For internalists the welcoming response is a response *to* a sensation, and the sensation is pleasure. Since the response is a response to pleasure and it is pleasure that is claimed to be good, this neatly satisfies the simple constraint. But for externalists the welcoming response is a *part of* pleasure. If you combine this view with the claim that pleasure is good, the view in question will not satisfy the simple constraint. The simple constraint claims that something can only be good when the subject has a positive response to *it*, and that would require a welcoming response (or some other positive response) *to* the bundle that already contains a welcoming response. In short, if we accept the simple constraint, then in order for externalist pleasure to count as good it would have to be true that we routinely welcome the thing that contains welcoming. This is implausible.

<sup>19</sup> Guy Fletcher raised this objection when this paper was presented at the 2014 Metaethics Conference. Namely, he pointed out that hedonism plus an externalist account of *pleasure* would violate SC3. He has also made a similar point in Fletcher (2013).

<sup>20</sup> Sumner (1996). It is unfortunate that so many different debates in philosophy are framed in terms of the labels internalism/externalism. Here the terms are currently being used in the service of a debate about *the nature of pleasure*—a debate entirely unrelated to the metaethical debate that is the main focus of the paper.

<sup>21</sup> Sumner (1996: 88).

<sup>22</sup> Heathwood (2007) is a contemporary defense of an externalist view of pleasure.

C5.P39 How problematic is this? First, it is important to note that it doesn't threaten my central claim which is this: at least part of the attraction of hedonism historically has come from the fact that theorists intuitively wanted to capture something like the simple constraint and thought that equating goodness facts with pleasure was a way to do this. Even if externalist theories are currently more popular with philosophers, historically many hedonists were internalists and others appear to have been confused or undecided about which view to accept.<sup>23</sup> Moreover, I suspect internalism about pleasure is the naïve view, in the sense that it is the one most of us begin with. But then, since internalism about pleasure is compatible with the simple constraint, my claim about the historical allure of hedonism may still be correct.

C5.P40 Second, the problems here may be pretty minor. Clearly it is the welcoming response (whether thought of as *part of* pleasure or as a reaction *to* pleasure) that underwrites many people's sense that pleasure is good for the one who experiences it (whether or not it is an all-things-considered good is a different matter). Given that this is so, an externalist about pleasure could make several moves. She could modify the simple constraint slightly. For example a hedonist who is externalist about pleasure could adopt SC4, which, in spirit, is not so different from the earlier versions of SC:

C5.P41 SC4: A necessary condition of X's being a basic, intrinsic prudential good for a subject, S, at T is that X be the kind of mental state that has a positive welcoming response built into it.<sup>24</sup>

C5.P42 Alternatively, the externalist could simply say that we misspeak when we say that pleasure is what has intrinsic value. Rather it is the central element of pleasure—the sensation at its core—that has value, and it has this in virtue of the fact that we naturally welcome it. Finally, one might simply adopt an internalist account of pleasure. After all, it does seem much more natural to say that we *welcome* pleasure (as opposed to saying that pleasure is partly a welcoming).<sup>25</sup>

C5.P43 Though I do not have space to consider other theories besides hedonism that focus on occurrent attitudes or experiences, it should be obvious that many of them will also have the resources to satisfy SC3.

<sup>23</sup> For a discussion of the different strands of thought running through the tradition see Sumner (1996: 87–91).

<sup>24</sup> A value pluralist who counts pleasure as one value among others could combine SC4 with the earlier SC3 to get a hybrid principle.

<sup>25</sup> For recent, more sophisticated internalist views of pleasure, see Crisp (2006), Smuts (2011).

C5.S5 5.2.1.2. *Desire Fulfillment Theories*

C5.P44 Although desire theories are different from hedonism, desire theorists also seem to be concerned with something like the simple constraint.<sup>26</sup> The evidence for this is indirect, but nonetheless powerful. Without trying to name all the different desire theories or go into their details, we can point to a particular direction that the development of such views has taken over time, one that is suggestive of concerns about positive response to realized basic intrinsic good.

C5.P45 The most basic form of desire theory claims that what is good for a person is for her actual desires to be satisfied. It is the desire satisfaction—a state of affairs in the world—that is the bearer of intrinsic value. The mental state ‘desire’ is simply relevant as the device that tells us which states of affairs benefit which people. Only those states of affairs that are the satisfaction of *my* desires benefit *me*.

C5.P46 What matters for our purposes is the way these views have developed over time. Most theorists take ‘desire’ to be a prospective attitude.<sup>27</sup> I desire now what I shall (perhaps) get later. It is this feature of desire that creates problems for basic desire theories. It is usually assumed that the satisfaction of a desire benefits a person at the time the desire is satisfied. However, a person might desire X at T, but when her desire is satisfied at T2 and she is aware of this, it might be that she is disappointed in X. The problem isn’t just that she is unhappy. Desire theorists pride themselves on being able to accommodate the fact that not everyone cares about happiness. Some individuals do quite well in life despite having little happiness. The issue is not happiness per se, but rather *lack of any kind of positive response* to the

<sup>26</sup> It is sometimes said that one of the most intuitive features of desire theories is that they (better than other theories) capture the internalist constraint. See e.g. Arneson (1999), Noggle (1999), Brink (2008), Dorsey (2017). However, what is meant is *not* that desire theories preserve something like the constraint I defend here. Rather such comments presuppose an entirely different vision of what prudential internalism amounts to, a vision I think we should reject (see my arguments against Rosati in Section 5.3). Although the features of desire theory commonly thought to preserve internalism are not my focus, I think one can trace in the development of desire theory an *implicit commitment* to something like my constraint. So far as I know, such a commitment has never been explicitly articulated by desire theorists or labeled by them as internalism. But that the commitment is there at all strikes me as highly significant and telling.

<sup>27</sup> Sumner (1996: 129). Heathwood (2006) bucks this trend, using ‘desire’ to name a concurrent attitude, a pro-attitude towards something at the time you have it. He thereby becomes immune to the particular problem for desire theories outlined here. However, it is doubtful that his theory (a species of what Parfit (1984) called “preference hedonism”) would be classified as a form of desire theory. Most simple desire theories do face the problem.

putative good. Not only does it not make her happy or give her pleasure, but now that she ‘has’ it she sees no value in it.

C5.P47 This is just the same point made earlier: if something is supposedly intrinsically good for a person at a time (and she is aware of this thing) we expect there to be some kind of positive response to the good thing at that time. Theorists can differ about the type of response that matters, but if there is no positive response at all we cannot sustain the thought that the thing really is good.

C5.P48 Because there are so many ways that prospective desires can be defective and thus so many ways that the satisfaction of prospective desires can disappoint us, simple desire theories have very few advocates. However, it is precisely the recognition of this problem *as a problem* that I want to highlight, because theorists with otherwise quite diverse approaches have shared this intuition. Desire theorists *themselves* have tried to adjust their view to avoid this problem, which suggests they accept that the counter-examples point to a real issue. This reflects an implicit acceptance of the simple constraint.

C5.P49 In response, desire theorists have moved to some type of informed desire theory. Some of these views are like the simple version in holding that the satisfaction of informed desires is of intrinsic value. But others depart from this somewhat. Although still called desire theories, here the idea is that under certain improved epistemic conditions a person would be able to act as her own best advisor.<sup>28</sup> From the better vantage point she would be able to ‘see’ the prudential goodness facts, facts about goods but also facts about choices and lives. The individual’s good is defined as what from this special epistemically advantaged position she would want for her actual self if she were to contemplate her actual self’s position as someone about to assume her actual self’s position.<sup>29</sup> The advantaged self might want particular intrinsic goods for her actual self, but she might also want her actual self to make certain choices that will lead (perhaps only in the long run) to intrinsic goods.

C5.P50 The problem with all such accounts is that it still seems possible that someone might get the thing her informed self wants for her and yet not find any value in it. Alternatively, she might make the *choice* that her informed self wants her to make and persevere in this direction, yet never come to respond positively to or see value in the things she thereby acquires. In the simpler case we find it implausible to suppose that a thing could be intrinsically good at a time when a person (who is aware of it) sees no good in it. In the slightly more complex case, we find it impossible to believe

<sup>28</sup> Railton (1986a; 1986b).

<sup>29</sup> Railton (1986b: 174).

that a *choice* is good, unless we are confident that it will at least *eventually* lead to encounters with things we respond positively to. But when we have no confidence that continued pursuit of a particular path will eventually lead to positive response of some form, we find it impossible to view the case as one where an individual is proceeding down a path towards something good for her. Even when the ideal epistemic conditions are described as incorporating full information and full rationality, philosophers continue to have doubts about whether such cases (cases in which the actual self finds no value in the things selected for her by the ideal self) will be eliminated.<sup>30</sup> In other words, the difference that remains between theorists who accept full information accounts of prudential good and theorists who don't does not appear to be a disagreement about whether a theory *should* satisfy SC3 but instead looks like a disagreement about whether that particular theory *in fact* satisfies SC3. Both parties seem committed to the idea that the theory would fail *if* it failed to satisfy the simple constraint.

c5.S6 5.2.1.3. *Objective List Theories*

c5.P51 Objective list theories are traditionally thought of as those theories that reject the necessity of any kind of psychological response. Although this means that most of them fail to satisfy SC3, this is hardly damning since most of them explicitly reject any form of existence internalism. However, it is worth noting that a number of theorists have recently tried to amend this very feature of objective list theories, recognizing it as the most problematic part of the view. As Shelly Kagan puts it, "Friends of an objective account of well-being seem forced to accept the unappealing claim that I could be extremely well-off, provided that I have the right objective goods in my life, even though these things hold no appeal for me, and I am, in fact, utterly miserable."<sup>31</sup> It is precisely the unattractiveness of such a claim that leads Kagan (as well as a number of other theorists interested in objective views)<sup>32</sup> to consider a modified version according to which objective goods can improve a person's life only if he or she has positive attitudes towards them. On Kagan's view a person must *enjoy* the objective goods. Of course it is not at all clear whether these theorists would be willing to distinguish between realized good and merely potential good and so not clear whether they would embrace the simple constraint. But the quote from Kagan certainly does seem to suggest the simple constraint. For it emphasizes the extreme implausibility of "having" or "possessing" an intrinsic good at the very time it makes you miserable.

<sup>30</sup> Sobel (1994), Rosati (1995). <sup>31</sup> Kagan (2009: 254).

<sup>32</sup> e.g., Adams (1999); Darwall (2002).

C5.S7

### 5.2.2. Further Objections

C5.P52 Despite its initial plausibility, many theorists continue to worry about certain putative counterexamples to SC3. These challenge the necessity of positive response (or positive engagement) in cases where the subject clearly has knowledge or awareness of the putative intrinsic good.

C5.P53 A common case is one where a person is in the process of acquiring some new skill or transitioning into some new set of circumstances. For example, consider Maria who, unlike Samia, has the talent to become an excellent writer and would thrive as one if she could just get through the excruciating first years. Surely (says the objector) it is good for Maria to keep going despite the fact that currently she has no positive response? Yes. However, the sense in which it is good for her to do this is instrumental. It is a good choice to persevere because genuine intrinsic goods lie in wait for her if she sticks with it and acquires the skill.

C5.P54 A different case raises more serious concerns. Consider someone in a severe clinical depression. He is so depressed he does not currently have the capacity to respond positively to anything. He feels no pleasure. He is not happy. He values nothing. His family is loving and supportive and is trying to get him into a good treatment program. Surely (says the objector) it is good for him to go into treatment? Surely the love and support of his family are good for him even now, while he cannot appreciate their efforts? The simple constraint seems to imply that nothing is intrinsically good for this person right now. But (so the objection goes) that can't be right.

C5.P55 First, it is important to notice that the case goes beyond clinical reality (and so may never actually be instantiated). People who are severely depressed often cannot feel *much* pleasure. And they often value *far less* than they did beforehand. But it is not clear that there really are people who remain self-aware and yet have *no* capacity to respond positively *to anything at all*.

C5.P56 Yet there are people whose capacities for positive response are severely limited. So the question remains, what is good for them? The answer depends on thinking about the structure of good. Certain kinds of claims involving 'good' are true of this individual.

C5.P57 If the person is treatable, then there exist possible continuations of his life in which he could realize many basic, intrinsic prudential goods that he cannot realize now. Assuming that some of these possible life continuations are ones that contain net positive intrinsic value, he has a normative prudential reason now to try and bring about one of these. Whether or not he realizes it, he has a reason now to try and end his depression. Mostly likely he doesn't realize it. But it is also true that those who love him have

reason now to help him out of depression and should do so precisely because he can't appreciate his own situation. My account does say that very few things count as *basic intrinsic prudential goods* for him *at this point in time*. But this seems right. Indeed, it is part of the explanation of the badness of depression that it narrows our evaluative connections to the world in this way. It undermines us as valuers and as responders.

- C5.P58 Some people, however, are still not satisfied. Surely, they argue, the love and support of the family is intrinsically good *now*? It is not clear to me why we should want to say this. *If* he currently has *no* ability to respond positively to support and cannot maintain positive engagement with those who love him, it seems that currently the love and support is quite literally not benefiting him. And because of this, it fails to be intrinsically good *for him now*. This is compatible with the claim that there are strong reasons now for providing love and support. It may lead to the individual getting the right help. The love and support may also be good in the sense that it is good to keep trying to engage someone who is depressed. You never know when your expressions of love might “get through.” Finally, some theorists might wish to say that loving another person is a (morally) good thing for the one who loves or that it is just non-relationally good—good *simpliciter*—that there be more love in the world. All of this is compatible with recognizing that the love and support is not intrinsically good *for* the depressed person while he is so deeply depressed.

C5.S8 5.3. MOTIVATION TO PURSUE THE GOOD?

- C5.P59 So far I have focused entirely on one type of constraint, a constraint on truths about *realized* good of the form:

C5.P60 “X is a basic, intrinsic prudential good for S now.”

- C5.P61 But what about *potential* good? What about truths of the form:

C5.P62 “Y *would be* good for S”?

- C5.P63 Should such truths motivate agents in the way that some philosophers think moral truths should motivate individuals to act rightly?

- C5.P64 We need several clarifications before we can answer. To begin with, the sentence “Y would be good for S” is ambiguous. ‘Y’ could refer to some putative intrinsic good or to a choice. Moreover, it could be true that something—perhaps a particular relationship—would be intrinsically good for a subject S, but still false that she ought to pursue it. For it might also be true that the relationship (which is a good) is part of a path through life that is not as good overall. Even if the relationship is



intrinsically good, if it is not part of one of the good continuations of S's life, then it is not a good choice. So if we want to consider issues of motivation, we should focus on true claims about good choices. If any prudential facts should motivate individuals to act, surely it would be these.

C5.P65 Are all subjects such that if they are minimally rational and confronted with such facts they will be (at least somewhat) motivated? The right answer, I think, is: no.<sup>33</sup> When it comes to merely potential good, we should be motivational externalists. To see why, we first need some clarity about what is at stake. We are interested in existence internalism, which is a *metaphysical* thesis. It is easy to confuse intuitions in this realm with intuitions about judgment internalism. Judgment internalism holds that there is a necessary connection between sincere judgment involving evaluative concepts and motivation. For all I know, judgment internalism may be true for prudential good. Perhaps the notion of prudential good is such that if you genuinely believe Y is the best choice for you, you will necessarily have some motivation to choose Y. I take no stand on that. But notice that if judgment internalism is true, people will be motivated by anything they believe to be good, whether or not it really is good.

C5.P66 However, existence internalism is concerned with the relation between normative facts naturalistically construed, on the one hand, and human responsiveness, on the other. Theorists who want to capture some form of existence internalism tend to equate prudential value facts with facts they think most people care about. Most people care about their happiness, even if that is not all they care about. Because of this some theorists have hoped to secure the right kind of internalist connection by equating prudential value facts with facts about what would make someone happy. Indeed, the constraint would work if we could be sure that everyone is always, at least to some degree, motivated by facts about what would make them happy. But this is not the case. Nor am I aware of any other candidate set of goodness facts that people are reliably motivated to pursue, independently of their evaluative beliefs about prudential good.

C5.P67 In short, to test existence internalism we have to imagine a subject who is considering a choice that, according to our favored theory, is good for her. She need not think of the choice or its outcome *as good*, but she must have correct descriptive beliefs about what her life would have in it if she made the choice and what her life would be like for her at various points in the

<sup>33</sup> Many theorists incline towards 'yes.' Rosati (1996) defends an indirect form of motivational internalism. Hall and Tiberius (2016) assume the truth of motivational internalism, and argue that one point in favor of subject-dependent theories is that they are well placed to explain this truth. However, it should now be clear that not all subject-dependent theories actually preserve motivational internalism about *potential good*.

future if she made it. The question is whether we can identify any plausible potential goodness facts that we can be certain she will be motivated by. There is no reason to think that we can.

C5.P68 If I am right, then we should be motivational externalists when it comes to *unrealized, merely potential* good. Of course, even externalists can allow that people are *sometimes* motivated by such facts. What they deny is that such motivational connections are necessary. And that seems right, particularly once we get clear about the structure of good. Recall that sometimes it can be good for a person to change her circumstances or herself dramatically. It is highly plausible that, at least once in a while, the best path forward through life for an individual will be one of the paths that requires dramatic self-change. But then our theory will say dramatic change is the best choice. And in these kinds of cases, even if we have great confidence in our theory, we simply cannot guarantee that such facts will have motivational force for the agents.

C5.P69 To help make clear that the problem is not specific to a particular theory, I will consider two examples. For the first I will assume the truth of hedonism, and for the second the truth of some form of value pluralism.

C5.P70 Consider Miserable Maud, a religious ascetic.<sup>34</sup> She believes pleasures come from the devil and must be avoided as much as possible. She lives her life in a way that pretty successfully minimizes pleasure. But Maud is certainly capable of experiencing pleasure and sometimes despite herself she experiences some. According to hedonism the best path forward for Maud is one that leads her to experience a great deal of pleasure. However, Maud is unlikely to embark on such a path without a dramatic change in her outlook, since her current attitudes lead her to avoid pleasurable things and also serve to dampen her experience of the few pleasures she unwittingly stumbles upon. Not only is the best life for Maud one in which she has more pleasure, the best life is one in which she changes herself in ways that make it possible for her to experience much more pleasure than her current attitudes allow. Thus if hedonism is true it may well be that the best choice for Maud now is to begin the process of altering her religious beliefs. She need not abandon religion. But she needs a religious view that (unlike her current one) encourages pleasure. The internalist question is whether we can expect Maud to be motivated *now* by knowledge of the descriptive facts about how much pleasure she could experience if only she changed her beliefs. Is it likely that she would be motivated to make such a choice? Absolutely not! From her current position she has zero motivation to change.

<sup>34</sup> The example is borrowed from Fletcher (2016: 67–8).

- c5.P71 Of course, for all I have said so far, the problem could be a problem with hedonism rather than motivational internalism about potential good. If hedonism is false, then it will not be surprising if someone is not motivated by knowledge of facts about how she could change to maximize her experience of pleasure. However, the problem arises for other views as well. To see this, consider the next example.
- c5.P72 Consider Miserable Martha. Martha is not always miserable, but she does suffer from regular bouts of serious unipolar depression, and when in the grips of these spells she is miserable. When she is not depressed she is quite artistic. She strongly suspects that part of what contributes to the quality of her art is something she gains from her depressive spells. Suppose she is right. And suppose she has the following option: she could take a medication that would relieve her of her depressions, but it would also alter slightly her personality. She would be capable of living a wonderful life as the person she would become, but she would not be nearly as creative and the quality of her artwork would be significantly less. She could, however, have healthy relationships that are currently impossible given her recurrent depressions. And she could succeed in many other, equally gratifying, though non-artistic pursuits. Here I assume a pluralistic view of welfare that allows for many different basic, intrinsic prudential goods, including creative production of art, other kinds of achievement, and relationships.
- c5.P73 Suppose the best possible continuation of Martha's life—the one with the greatest net positive intrinsic prudential value—is one of the lives she could lead only with the medication. This means she has most reason right now to give up her artistic pursuits, take the medication, and embark on a highly different path through life. Should we assume that the person she is now—during one of her non-depressed moments—would be motivated by knowledge of these facts? Again, recall that what she knows are just the descriptive facts about what her life will have in it and what it will be like for her once she becomes the post-medication person. Will knowledge of such facts *necessarily* motivate her to choose the medication? It seems clear to me that we cannot expect this. Indeed, it is quite likely that the facts won't motivate her at all, particularly if from her current position she can't personally identify with the individual she would become if she pursued that path.
- c5.P74 I don't think this is a problem. We shouldn't really expect her to be motivated by facts about her potential good that are so removed from her current circumstances. The prudential goodness facts sometimes point in directions that, from our limited, current perspectives, seem odd. It may just be that the complexity of future prudential value ensures that we cannot always care about our future good from afar. Motivational externalism about potential good begins to seem like the sensible option.

### 5.3.1. Two-Tier Internalism?

C5.S9

Before we settle for SC3, we should examine carefully the view of Connie Rosati, labeled “two-tier internalism.”<sup>35</sup> To date this is the most well-developed alternative to my view.

C5.P76

To grasp two-tier internalism we need a couple of definitions. First, we need the notion of “ordinary optimal conditions.” These are just the conditions of judgment most of us occupy most of the time. They exclude the obvious sorts of conditions that distort judgment (or make it impossible)—things like being asleep, being drunk, being delusional, being in a state of intense emotion, or being clinically depressed.<sup>36</sup>

C5.P77

Second we need the notion of “ideally authoritative conditions.” This is the idea of conditions that improve judgment to the point where it can’t go wrong and which are authoritative for that reason. Most of us realize our ordinary judgments are hampered by various limitations. We lack knowledge. We are prone to many forms of irrationality. Thus, it is natural to suppose that if we had full knowledge, or could be always rational, our judgments would presumably be much better. If we push such ideas to the limit we arrive at the notion of ideally authoritative conditions—the set of epistemic conditions that are ideally authoritative for making judgments about prudential value.<sup>37</sup> I shall refer to these conditions as C\*.

C5.P78

We can now explain two-tier internalism as follows. Consider a case of *potential* good (as we shall see it works the same way for realized good). Facts about potential good can typically be translated into claims about good choices in the present. So, suppose it is true that Y is the best choice for S now.

C5.P79

Speaking roughly, two-tier internalism requires first that S herself under ideally authoritative conditions recognize this fact about her good (that Y is the best choice). For various reasons the actual requirement needs to be formulated in a slightly more complicated way as follows: it must be true that in C\* if S contemplates the situation of her actual self *as someone about to assume the actual self’s position*, she would want her actual self to choose Y.<sup>38</sup>

<sup>35</sup> Rosati (1996: 307). Alexander Sarch (2011) argues against Rosati’s two-tier internalism. However, he does not contemplate the possibility of an internalist constraint on realized good like the one defended here.

<sup>36</sup> Rosati (1996: 304–5). <sup>37</sup> Rosati (1996: 304).

<sup>38</sup> For ease of exposition, I often leave out these details, but they should always be understood to be present. I sometimes use short phrases like “what S would want in C\*” to stand in for longer phrases of the sort: “what S would want in C\* for her actual self if she were contemplating the circumstances of her actual self as someone about to assume the actual self’s position.”

C5.P80 The second requirement of two-tier internalism is that under ordinary optimal conditions a person, S, cares about what she would want in C\*. This means that even if in ordinary circumstances she can't recognize facts about her good *as facts about her good* (and so can't be motivated by them), she still cares at that time about *what her ideal self would want in C\**. Of course, presently she can't make the connection between what is really good for her (Y) and what the ideal self would want in C\*. But the fact that right now she cares about what her ideal self would want in C\* means that *if* she could make the connection—if she could come to see that Y is what the ideal self would want in C\*, she would right now have *some* motivation to pursue Y.

C5.P81 Rosati clearly intends two-tier internalism to serve as a constraint on both facts about *realized* prudential good and facts about *potential* prudential good. In other words, she does not make the modal distinction that has been central to my discussion so far. For that reason, let us briefly consider in turn, two-tier internalism as a constraint on *realized* good, and then as a constraint on *possible* good.

C5.P82 Rosati and I agree that there is an internalist constraint on facts about current realized intrinsic good—that there must be some kind of psychological link between a subject and facts about what is good for her now. We simply offer two different accounts of what that constraint is. Like SC3, Rosati's constraint is best understood as a necessary condition only. Although two-tier internalism and SC3 are naturally understood as competing accounts of a single constraint on facts about realized good, nothing rules out in advance the possibility that there are two constraints.

C5.P83 This turns out to be relevant in the following way. It is incredibly difficult to undermine the two-tier internalist requirement for realized good *directly*. To do so we would need to have a clear case in which we could all agree that something, X, is currently good for a subject, S. And we would then need to establish in that case two-tier internalism fails to hold. That would decisively show that the two-tier constraint is not necessary, as claimed. However, for various reasons, it is incredibly difficult to know in any given case whether or not two-tier internalism is satisfied. So rather than undermine it directly, my aim is to show its irrelevance.

C5.P84 First, it seems to me that no matter what the status of two-tier internalism, we still need SC3 in the case of realized goodness facts. Suppose, for example, a theory were to say that writing is intrinsically good for Samia at the very time when she is so utterly indifferent to it. I confess, I simply can't imagine anything that would convince me to view something as *intrinsically* good for her at that time given the psychological facts about her at that time. So SC3 remains necessary. Second, it seems important to note that one of the prime motivations for two-tier internalism fails in the case of realized goodness. Two-tier internalism was formulated in response to worries about

the ability of a subject to recognize and positively respond to facts about her own good. It is designed to provide a kind of *indirect psychological link* in cases where a *direct link* is lacking. But once we make the modal distinction, and recognize the importance of SC3 for facts about realized goodness, we also see that indirect links are not needed, because direct links (currently existing links) are required. SC3 rejects any theory where they are absent. In other words, once we focus narrowly on facts about realized good, we see that there are no cases in which X is currently good for S but S as she actually is fails to respond positively to X in any way. But since no such facts exist, there is no need to salvage internalism for these cases by forging an indirect link through facts about what S would want in C\*.

C5.P85 Nonetheless, one might think that we need something like the two-tier requirement when we focus on facts about possible good. It is unlikely that any theory of welfare will be able to ensure that agents are always reliably motivated to pursue what would be best for them. But one might hope that two-tier internalism could forge an *indirect link* between the facts about potential individual good and the subject's current motivations.

C5.P86 Consider again Miserable Martha for whom (let us assume) the best choice is medication and self-transformation. If the two-tier requirement is satisfied, this means both (a) that in C\* Martha would want her actual self to adopt the life that medication makes possible, and (b) right now (assuming she is not currently in the grips of depression) Martha cares about what her ideal self would want in C\*. The question is whether the best theory of welfare can ensure both that Martha in C\* wants the right things and that actual Martha cares about the desires of her ideal self. To put the point more dramatically, should we *reject* a theory of welfare if it fails to ensure these two claims? I think not.

C5.P87 There are several reasons for this, though I grant that the first reason is not widely shared. I am skeptical of the idea that we can describe a set of ideal epistemic conditions that could be occupied by a real human being *and* would reliably reveal to this human occupant all the truths about her good.<sup>39</sup> In other words even if we make sense of the idea of an agent who is *fully* informed and *fully* rational, I have no faith that the pronouncements of such an idealized self would really track the goodness facts. I am not sure we should trust the pronouncements of such a self. Nor do I think there is some way to amend the description of ideal conditions to fix the problem. This is the basis for my skepticism about full information theories of prudential good. But it is relevant to the question of whether to accept two-tier internalism, for although it is not a theory of prudential value, it is a

<sup>39</sup> Similar skepticism is expressed by Sobel (1994) and Rosati (1995).

proposed constraint on such theories. If there is no  $C^*$  such that the pronouncements from  $C^*$  are always correct, then we cannot require a theory of welfare to conform to a restriction that appeals to  $C^*$ .

C5.P88 However, even if one does not share my skepticism about  $C^*$ , it is possible to raise worries about two-tier internalism. Two-tier internalism assumes that there is some description of ideal conditions  $C^*$  such that  $C^*$  satisfies two requirements. On the one hand,  $C^*$  is such that all welfare agents view  $C^*$  as authoritative (this is what underwrites the confidence that under ordinary optimal conditions agents will care about what their ideal selves want in  $C^*$ ). And on the other hand,  $C^*$  is such that agents in  $C^*$  reliably perceive the goodness facts (or at least reliably want for their non-ideal selves the things that are really good). The problem, however, is that there may be no  $C^*$  that can satisfy both conditions.

C5.P89 Suppose for the sake of argument that there is some set of ideal epistemic conditions  $C^*$  such that an agent in these conditions reliably judges her own good. Still, actual agents do not all agree about what these conditions are. Probably it is true for the vast majority of people that there is *some description or other* of better epistemic conditions they would accept as authoritative. Let us call the conditions accepted by Martha  $C^{+Martha}$ , and the ones accepted by Maud  $C^{+Maud}$ , and for each person the conditions she would view as authoritative are  $C^{+hername}$ . In ordinary optimal conditions Martha cares what she would want in  $C^{+Martha}$  and Maud cares what she would want in  $C^{+Maud}$ . But unless  $C^{+Martha}$  is the same as  $C^*$  this will not forge a motivational link with the true goodness facts. In short, if people do not generally agree about ideally authoritative conditions, then even if there is a  $C^*$  that offers a reliable epistemic perspective on the prudential goodness facts, it will be impossible to establish an indirect motivational link from individuals to those facts.

C5.P90 It is important to remember that many people, even when they are free of violent emotion, are not delusional, and so on, are willing to grant authority to less than ideal judgments—judgments made without full information or that are irrational in one or another sense. Nor can they all easily be brought to see the error of their epistemic ways. Even if very many people can, we have to remember that internalism is supposedly *a necessary condition*. My point is just that there is no value for  $C^*$  such that  $C^*$  offers the right kind of perspective on the goodness facts *and*  $C^*$  is viewed as authoritative by all people in ordinary optimal conditions. In the very best case the link remains highly contingent. Therefore we should not require a theory of welfare to satisfy the two-tier internalism requirement.

C5.P91 Accepting motivational externalism about potential good might be more distressing if we were also externalists about realized good. To embrace a dual form of externalism would be to accept that the psychology of

individual subjects places *no necessary constraint on prudential goodness facts*. That does seem deeply problematic. Since most theorists until now have not distinguished realized good and potential good and so have not considered whether the constraints for each might be different, it is easy to see why a claim like two-tier internalism would be popular. But after reflection it seems that it just isn't necessary. The right link between the psychology of agents and goodness facts is a link that holds when *basic, intrinsic goodness is realized*. And once we see that, we see we have all we need.

C5.S10

#### 5.4. CONCLUSION

C5.P92 We have covered a lot of ground. I first reviewed my understanding of the “structure of good,” emphasizing a modal distinction between realized intrinsic prudential goods and merely potential good.

C5.P93 Against this backdrop I argued that we should understand existence internalism for prudential value as a constraint on realized intrinsic good. In an attempt—at least for now—to keep the formulation of the constraint broad so that many theories might satisfy it, I proposed the following formulation:

C5.P94 SC3: A necessary condition on something, X, being a basic, intrinsic prudential good for a subject, S, at a time T, is that:

C5.P95 If S is aware of X at T, S must either have a positive response to X at T, or S must have an on-going positive engagement with X that still holds at T, *or*

C5.P96 If S is unaware of X at T, it must be true that S would have had a positive response to X at T if she had been aware of X at T.

C5.P97 This is a merely necessary constraint. Positive response should be taken to include a wide range of positive psychological states.

C5.P98 Having defended SC3, I considered whether we should also want some form of internalism constraint that would guarantee that subjects are motivated, at least to some degree, by truths about their own potential good. We should not. We should accept motivational *externalism* about potential good. Individuals will not necessarily be motivated to seek their good, but *if* they seek it, and get it, they will respond positively to it.

C5.P99 Finally, I considered and rejected two-tier internalism, a highly sophisticated attempt to forge an indirect link between prudential goodness facts and an individual's motivations. SC3 is superior to two-tier internalism when we focus on facts about realized goodness. And in the case of potential good it is ultimately wiser to accept that no internalist constraint, even an indirect one, can be had.



C5.S11

## References

- Adams, Robert Merrihew. 1999. *Finite and Infinite Goods*. Oxford: Oxford University Press.
- Arneson, Richard. 1999. "Human Flourishing versus Desire Satisfaction," *Social Policy and Philosophy* 16 (1): 113–42.
- Brink, David O. 1989. *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press.
- Brink, David O. 2008. "The Significance of Desire," in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 3*. Oxford: Oxford University Press, 5–45.
- Crisp, Roger. 2006. "Hedonism Reconsidered," *Philosophy and Phenomenological Research* 73 (3): 619–45.
- Darwall, Stephen. 1983. *Impartial Reason*. Ithaca, NY: Cornell University Press.
- Darwall, Stephen. 2002. *Welfare and Rational Care*. Princeton, NJ: Princeton University Press.
- Dorsey, Dale. 2015. "The Significance of a Life's Shape," *Ethics* 125 (2): 303–30.
- Dorsey, Dale. 2017. "Why Should Welfare 'Fit'?" *Philosophical Quarterly* 67 (269): 685–708.
- Feldman, Fred. 2000. "Basic Intrinsic Value," *Philosophical Studies* 99 (3): 319–46.
- Fletcher, Guy. 2013. "A Fresh Start for the Objective List Theory of Well-Being," *Utilitas* 25 (2): 206–20.
- Fletcher, Guy. 2016. *The Philosophy of Well-Being: An Introduction*. New York: Routledge.
- Glasgow, Joshua. 2013. "The Shape of a Life and the Value of Loss and Gain," *Philosophical Studies* 162 (3): 665–82.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Oxford University Press.
- Hall, Alicia and Valerie Tiberius. 2016. "Well-Being and Subject Dependence," in Guy Fletcher (ed.), *The Routledge Handbook of Philosophy of Well-Being*. New York: Routledge, 175–86.
- Hawkins, Jennifer. 2014. "Well-Being, Time, and Dementia," *Ethics* 124 (3): 507–42.
- Heathwood, Chris. 2006. "Desire Satisfactionism and Hedonism," *Philosophical Studies* 128 (3): 539–63.
- Heathwood, Chris. 2007. "The Reduction of Sensory Pleasure to Desire," *Philosophical Studies* 133 (3): 23–44.
- Heathwood, Chris. ms. "What Is the Resonance Constraint?"
- Kagan, Shelly. 1998. "Rethinking Intrinsic Value," *The Journal of Ethics* 2 (4): 277–97.
- Kagan, Shelly. 2009. "Well-Being As Enjoying the Good," *Philosophical Perspectives* 23 (1): 253–72.
- Korsgaard, Christine. 1983. "Two Distinctions in Goodness," *The Philosophical Review* 92 (2): 169–95.
- Moore, G. E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.

- Nagel, Thomas. 1970. "Death," *Noûs* 4 (1): 73–80.
- Noggle, Robert. 1999. "Integrity, the Self, and Desire-Based Accounts of the Good," *Philosophical Studies* 96 (3): 303–31.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Railton, Peter. 1986a. "Facts and Values," *Philosophical Topics* 14 (2): 5–31.
- Railton, Peter. 1986b. "Moral Realism," *The Philosophical Review* 95 (2): 163–207.
- Rosati, Connie. 1995. "Persons, Perspectives and Full Information Accounts of the Good," *Ethics* 105 (2): 296–325.
- Rosati, Connie. 1996. "Internalism and the Good for a Person," *Ethics*, 106 (2): 297–326.
- Rosati, Connie. 2016. "Moral Motivation," *The Stanford Encyclopedia of Philosophy* (Winter 2016 edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2016/entries/moral-motivation/>.
- Sarch, Alexander. 2011. "Internalism About a Person's Good: Don't Believe It," *Philosophical Studies* 154 (2): 161–84.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press.
- Sidgwick, Henry. 1907. *The Methods of Ethics*, 7th edn. New York: Hackett.
- Slothe, Michael. 1984. "Goods and Lives," in *Goods and Virtues*. Oxford: Oxford University Press, 9–37.
- Smith, Michael. 1994. *The Moral Problem*. Malden, MA: Blackwell.
- Smuts, Aaron. 2011. "The Feels Good Theory of Pleasure," *Philosophical Studies* 155 (2): 241–65.
- Sobel, David. 1994. "Full Information Accounts of Well-Being," *Ethics* 104 (4): 784–810.
- Sumner, L. W. 1996. *Welfare, Happiness and Ethics*. Oxford: Clarendon Press.
- Temkin, Larry. 2012. *Rethinking the Good*. Oxford: Oxford University Press.
- Van Roojen, Mark. 2015. *Metaethics: A Contemporary Introduction*. New York: Routledge.
- Velleman, J. David. 1991. "Well-Being and Time," *Pacific Philosophical Quarterly* 72 (1): 48–77.
- Velleman, J. David. 2000. "Is Motivation Internal to Value? in *The Possibility of Practical Reason*. Oxford: Clarendon Press, 85–98.
- Zimmerman, Michael J. 2015. "Intrinsic vs. Extrinsic Value," *The Stanford Encyclopedia of Philosophy* (Spring 2015 edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2015/entries/value-intrinsic-extrinsic/>.