Well-Being, Autonomy, and the Horizon Problem

JENNIFER S. HAWKINS
University of Toronto

Desire satisfaction theorists and attitudinal-happiness theorists of well-being are committed to correcting the psychological attitudes upon which their theories are built. However, it is not often recognized that some of the attitudes in need of correction are *evaluative* attitudes. Moreover, it is hard to know how to correct for poor evaluative attitudes in ways that respect the traditional commitment to the authority of the individual subject's evaluative perspective. L. W. Sumner has proposed an autonomy-as-authenticity requirement to perform this task, but this article argues that it cannot do the job. Sumner's proposal focuses on the social origins of our values and overlooks the deep psychological roots of other evaluative attitudes that are just as problematic for welfare. If subjective theories of welfare are to be at all plausible they may need to abandon or modify their traditional commitment to the authority of the individual subject.

I. THE HORIZON PROBLEM AND SUBJECTIVE THEORIES OF WELL-BEING

Many individuals are strikingly limited in what they can appreciate about their lives. I shall refer to deficiencies in a person's ability to see the circumstances of her life as they really are as deficiencies of *horizon*. The image of the horizon seems apt, since literal horizons set the limits of what we can see and, in the types of cases I discuss here, both external environmental barriers and internal psychological barriers place limits on what individuals can 'see' or appreciate about their own lives. In addition to being troubling in their own right, limited horizons pose a serious problem for subjective theories of well-being, welfare or personal good. (I treat these terms as synonymous.) In what follows, I shall refer to this as *the horizon problem*.

The dominant trend in the theory of well-being is towards subjective theories. As I shall use the term, 'subjective' theories of well-being are those theories that make the determination of a person's good fully depend on the actual or hypothetical attitudes or sensations of the individual herself. These theories are popular for a number of reasons.

First, subjective theories fit well with the common view that well-being is highly individually variable and with the equally common intuition that many of the determinants of well-being are either psychological in nature or dependent on the psychological activities of the individual. For example, although few philosophers are hedonists these days, most recognize that positive quality of experience is a

doi:10.1017/S0953820808002963
Utilitas Vol. 20, No. 2, July 2008
Printed in the United Kingdom

significant contributor to well-being. Yet positive quality of experience is an entirely psychological phenomenon. Further, most people think that having significant values and commitments that shape and give meaning to one's life is necessary for well-being. But although the things valued are typically not psychological, valuing is a psychological activity of particular agents.

Second, many find it hard to believe that anything lacking the appropriate link to our psychology (actual or hypothetical) could really turn out to be good for us. Since subjective theories build an account of well-being directly out of highly individualized psychological materials, they seem well positioned to capture these intuitions.

While subjective theories come in many varieties, the most plausible and currently popular versions – desire theories and attitudinal-happiness theories – focus exclusively on attitudes of the subject (as opposed to simple sensations, such as pleasure and pain). By attitudinal-happiness theories I mean those theories that define happiness as an attitude as opposed to a mood, feeling or other pleasant sensation. Not all such theories equate well-being with happiness. For example, L. W. Sumner makes attitudinal happiness a necessary but not sufficient condition of well-being.[1]

Not only do they focus on attitudes, but these theories also share a core commitment to treating the evaluative perspective of the individual herself as authoritative regarding what is good for her. Since the focus on attitudes and the commitment to the authority of the subject's perspective are not necessary features of subjective theories *per se*, I shall refer to the sub-set of subjective theories that share these features as *subjective-attitudinal theories*, or 'attitudinal theories' for short.

Despite their appeal, subjective-attitudinal theories face significant problems in connection with deficient horizons. The first problem is that people's actual attitudes are often deeply prudentially flawed. Hence, any theory of well-being built exclusively upon these attitudes would be implausible. An attitudinal theorist of well-being therefore needs to build her theory around the 'correct' or 'properly authoritative' attitudes and this means she inherits the task of sorting flawed from non-flawed attitudes. The second problem is that this sorting task is not only difficult in its own right, but especially difficult to do consistently with the subjective theorist's commitment to granting ultimate authority to the evaluative perspective of the individual. Of course, some theorists will view the horizon problem as simply a problem about lack of information and the sorting task as simply the task of distinguishing informed from uninformed attitudes. However, as I will

[1] L. W. Sumner, *Welfare, Happiness, and Ethics* (Oxford, 1996), ch. 6. See also Fred Feldman, *Pleasure and the Good Life* (New York, 2004), ch. 4.

argue, there is much more to the horizon problem than merely a lack of information. If attitudinal theories are the best forms of subjectivism, and if my arguments are correct, then the whole subjective approach to theories of well-being faces a significant challenge.

My aim in this article is to bring significant, oft-neglected aspects of the horizon problem to light and to show why the horizon problem cannot be solved compatibly with the central commitments of subjective-attitudinal theories, as they are currently conceived. I begin by arguing that there is an irreducibly *evaluative* aspect to the horizon problem. I then examine in some detail L. W. Sumner's autonomy-as-authenticity requirement, which is a recent, positive proposal for solving the horizon problem construed as a partly evaluative problem. Sumner's proposed resolution incorporates some form of a procedural account of autonomy, along the lines of the one offered by John Christman.[2] Hence, as part of evaluating Sumner I evaluate the usefulness of an approach like Christman's *for a theory of welfare*. I argue that Sumner's proposal fails to resolve horizon issues even for the specific set of cases he considers; and, even more importantly, that the problem is broader and more complex than Sumner's diagnosis assumes. Many important cases of limited horizons have deep psychological roots – a fact that many subjective-attitudinal theorists appear not to have fully appreciated.

Indeed, as I shall try to show, reflection on the complex psychological dimensions of horizon cases suggests that no solution to the problem can satisfy the constraints adopted by subjective-attitudinal theorists. Of course, impossibility claims are notoriously difficult to establish in philosophy. So I will be content if I leave the reader sharing my impression that the prospects for such a resolution of the horizon problem are dim.

II. NARROW HORIZONS AND THE ATTEMPT TO DEFINE IDEAL CONDITIONS OF JUDGMENT

The problems for an attitudinal theory of well-being that arise from deficient horizons are best grasped through illustrations. So let us consider a couple of cases. I shall begin with Celie, the main character from Alice Walker's novel, *The Color Purple*.[3] What follows is my own synopsis of Celie's story:

Celie's early life is a study in trauma and abuse. From a young age, she is sexually abused by her stepfather (whom she falsely believes to be her biological

[2] John Christman, 'Autonomy and Personal History', *Canadian Journal of Philosophy* 21 (1991), pp. 1–24.

[3] Alice Walker, *The Color Purple* (New York, 1982).

father). Celie's mother dies while she is very young, and she is left alone in the care of this man. While still a child herself she is forced to bear him two children, who are then taken from her at birth (by the stepfather) and given to strangers. Eventually, Celie escapes this tormentor for a new one – the man her stepfather forces her to marry and who is simply referred to throughout the novel as Mr. —. Mr. — treats Celie like a slave, demanding not only that she care for his unruly children from a previous marriage and maintain a high standard of housekeeping, but that she work in his fields as well. He beats her mercilessly and constantly tells her how stupid and ugly he thinks she is.

Grim as this part of the story is, the novel as a whole is essentially a transformation story. Celie puts up with this abuse for years, simply accepting it as her lot in life. But very slowly a change comes about, largely as a result of her interactions with two strong-willed women who do not put up with abusive treatment in their own lives. The first is Sophia, the head-strong wife of Mr. —'s grown son Harpo. The second woman is ironically Mr. —'s mistress, the beautiful and talented Shug Avery, who eventually becomes more attached to Celie than to Mr. —. Celie's relationship with Shug completes a transformation that began long before, when she first watched Sophia stand up to Harpo. After years of nearly unimaginable abuse, Celie does what it was always in her power to do. She leaves Mr. —.

The philosophically interesting features of Celie's case are the questions it raises about her attitudes to her life. If we focus on the early part of the story, it seems safe to say that Celie's life is quite bleak and that it would be much better for her to leave Mr. —. While she eventually does this, we may wonder why she fails to see the value of this option *earlier*. For a number of reasons, Celie is clearly not, at first, able to see her situation in the way that we, the readers, are able to see it. The interesting question is what exactly Celie would need in order to see her life for what it is. For simplicity, in what follows I shall simply refer to the attitudes (actual or hypothetical) upon which subjective-attitudinal theories are built as *prudential judgments*. Admittedly, this is a bit linguistically strained, since the attitudes in question are not in all cases attitudes the individual herself would label 'prudential'. However, this does not affect the argument. Since these attitudes are linked by the role they play in defining prudential value, I find it easiest to refer to them this way. We can thus say that the interesting questions about Celie's case are questions about her prudential judgments and about what it would take to make her actual judgments better.

My second case is drawn from E. Annie Proulx's *The Shipping News*.[4] What follows is my synopsis of the story of the character Quoyle:

Quoyle, we are told, is an ugly, awkward, oversized boy from an unloving and at times abusive family that has made him feel an outcast all his life. As a child, Quoyle dreams that he has been accidentally placed in this family and

[4] E. Annie Proulx, *The Shipping News* (New York, 1993).

that one day his true family – composed of kind, good people – will arrive to claim him. Lacking any special talents and filled with self-loathing, the young adult Quoyle stumbles through life without close friends or support, moving from one low-paying menial job to the next. Then Quoyle – who has longed for love, but never had a relationship with a woman – meets Petal Bear.

Petal Bear is bad news through and through. She seduces Quoyle, but not out of any genuine interest in *him*. Seducing men is just what Petal Bear does for entertainment. When she is not seducing men, she manipulates them to get money, jewelry, clothing and other things she wants from them. She perceives Quoyle as an easy target. Eventually she marries Quoyle, presumably because she recognizes that he is too timid to oppose her or keep her from doing as she likes with other men and recognizes that he may serve as a useful backup source of cash and occasional shelter. Petal Bear is ferociously abusive of Quoyle. Despite the fact that she ends up having two daughters with him, she cares nothing for him or her children. She continues to steal money from Quoyle, to heap verbal abuse on him, and to take off for weeks at a time with other lovers making no attempt to hide what she is doing. Quoyle is rescued from her, literally, by accident. When she finally decides to leave Quoyle for good with one of her lovers, she and her partner are killed in an automobile wreck.

Petal Bear was certainly not mistaken in her estimation of Quoyle's vulnerability. Throughout this relationship, he continues to hope that she will change and grow to love him and the children. He continues to adore her. Initially he cannot stop mourning her death, even though he realizes fully that she died abandoning her children and leaving him for another man. Quoyle's transformation comes about only slowly. It begins when he is persuaded by a kind and charismatic aunt (whom he has only recently met) to move with her to Newfoundland where his father's family was originally from. In his new relationship with the aunt, he discovers a different kind of family life; and in his new surroundings, he finds a better, more rewarding niche for himself as the writer of a column for the small local newspaper. As he discovers something that he is good at, and develops a real set of friends, Quoyle manages to move away from the self-destructive patterns that defined his previous life.

Once again, for the purposes of the philosophy of well-being, the interesting questions are about what, initially, Quoyle can and cannot see about his own life and what this, in turn, reveals about common defects of prudential judgment. Why does he tolerate someone like Petal Bear? As with Celie, it is plausible to think that Quoyle faces several types of perceptual barriers. What is philosophically interesting is how exactly we should describe these limitations and how to adjust for them in a theory of well-being.

It should be fairly easy to see why Celie and Quoyle would pose problems for a desire theorist committed to building a theory of well-being around *actual* desires. It is simply implausible that satisfying Celie's actual desire to stay with Mr. — or Quoyle's desire to spend more time with Petal Bear would contribute to their well-being. Hence, any minimally plausible version of a desire theory has to appeal to the hypothetical desires that Celie or Quoyle would have under conditions

that are better for prudential judgment. Following a suggestion of Peter Railton's, the desire theorist will most likely equate Quoyle's well-being with what he would want for his actual self as he is now, if he were able to contemplate his current situation under better circumstances.[5] The question is precisely what these 'better' circumstances are.

It is perhaps less obvious why Celie and Quoyle pose a problem for attitudinal-happiness theorists. Sumner, whom I take as my illustration, argues that we should think of happiness not as a transient mood, but as a stable, positive attitude towards the conditions of one's life. This attitude is best thought of on the model of *life satisfaction*;[6] and this satisfaction, in turn, is understood to have 'both cognitive and affective components'.[7] Happiness requires a 'positive evaluation of the conditions of [one's] life' – a 'judgment that, at least on balance, it measures up favorably against your standards or expectations'.[8] But it also has an *affective* dimension. Sumner describes this as 'having a sense of well-being'. It is to find 'your life enriching or rewarding, or to [feel] satisfied or fulfilled by it'.[9]

Just as a desire theorist faces a problem because an individual can actually desire something that (intuitively) will not contribute to her well-being, Sumner's theory inherits the different but equally tricky problem that sometimes individuals are actually satisfied with their lives when intuitively they should *not* be. Some people are satisfied with lives a theory of well-being should not count as good. Moreover, since a theory of well-being is presumably meant not only to identify good lives, but also to arrange lives on a scale from best to worst, Sumner inherits the problem that even people who clearly are not 'satisfied' may still rank their lives as being far better than they are.

The cognitive aspect of happiness requires that the individual judge her life to be, overall, good and living up to her expectations. But an individual's assessment of her life will, of course, depend largely on what she knows about other lives, what she sees as possible for her, and importantly, on what she values. An individual with little sense of how things might be different and with low expectations of what life has to offer may assess her own life as being pretty good. But this may be in sharp contrast to the assessment of a third-party observer, aware of all the other opportunities unrealized in her life.

Similar distortions can enter at the affective level. For Sumner, not only must one judge the conditions of one's life positively, but one must

[5] Peter Railton, 'Moral Realism', *Philosophical Review* 95 (1986), p. 174; See also, Peter Railton, 'Facts and Values', *Philosophical Topics* 14 (1986), pp. 5–31.
[6] Sumner, *Welfare*, p. 149.
[7] Sumner, *Welfare*, p. 145.
[8] Sumner, *Welfare*, p. 145.
[9] Sumner, *Welfare*, p. 146.

also have a *felt sense* of well-being – a *feeling* of being fulfilled. While Sumner does not himself elaborate on the nature of affective states, it seems plausible to view them like other emotional states. Feelings will then have an interesting dual nature: both a phenomenological aspect and a judgmental aspect.[10] The assessments implicit in them may not be as well articulated as cognitive judgments, and they may occupy a place in the mind that is semi-independent from the part that makes the cognitive judgments. Nonetheless, they are significant for the theory of well-being. They reflect the subject's sense that her current feeling state is better than, or at least as good as, other states she has been in and that it is as good as can be expected. Since emotional or affective states contain assessments, distortions may enter in the same way. Just as limited experience can distort cognitive assessments of the general conditions of one's life, so too can it distort one's affective sense of how well things are going and of what kind of existence is acceptable.

We are now in a position to see why Celie and Quoyle pose a problem not only for desire theorists, but for attitudinal-happiness theorists as well. Clearly we cannot say of Celie that she is 'satisfied' with her life. She certainly wants things to be different. For example, throughout her years of living with Mr. —, she desperately wants not to be beaten. But the problem is that Celie seems not fully to grasp how much better things could be for her. Part of Celie's problem seems to be that, as a product of years of having Mr. — tell her how stupid and clumsy he thinks she is, she has come to believe it. Hence, her sense of just how bad her life is is gravely distorted by her sense of what someone like her (talentless and stupid) can expect in life. Although we need not imagine that she enjoys her life with Mr. —, she likely assesses it as being much better than we think it is. If so, an attitudinal-happiness theorist needs to correct for this.

Attitudinal-happiness theorists make the needed corrections in two ways. First, they refuse simply to equate well-being with professed happiness or satisfaction and, second, they place extra conditions on the judgments that count. For example, Sumner views professed happiness or satisfaction as a necessary (but by no means sufficient) condition for well-being; and he only counts professed happiness as a genuine indicator of well-being if the judgment is *well-grounded*. My judgment that my life is going well counts as informed and hence as well-grounded

[10] That affective states have a judgmental component is now a widely accepted view. By judgmental component, I do not intend to commit myself to a cognitivist view of emotion *per se*, but merely to some account that develops a strong link between emotional response and assessments. Appraisal theories, cognitivist theories, and perceptual theories of emotion all fit this bill. For a thorough overview of the literature on emotion, see Ronald de Sousa, 'Emotion', *The Stanford Encyclopedia of Philosophy* (http://plato.stanford.edu/entries/emotion, 2003).

if other information would not alter my assessment of my life were I to possess it.[11] Just as we can ask under what conditions our desires actually track what is good for us, so we can ask when self-assessments of life satisfaction are actually trustworthy. However, this raises again, albeit in a slightly different form, the question of what it takes to correct prudential attitudes.

At this point, it is also important to highlight the constraints under which subjective-attitudinal theorists operate when trying to answer this question. There are really two constraints, operating at different levels. But they are connected by a core commitment to the authority of the subject's perspective. At the first level, a subjective theorist of well-being is committed to remaining neutral about which specific values contribute to a good life. This is required by the idea that individual welfare values should emerge from the subject's own evaluative perspective, rather than being stipulated by the theorist.

However, as we have already seen, the attitudinal theorist cannot avoid engaging in a sorting task, which already is a way of specifying the kinds of values that can contribute to a good life – albeit indirectly, by specifying the mental attitudes and mental profiles of those who count as good judges. Hence, a second constraint enters when we turn to consider which sorting mechanisms are legitimate. If the sorting task itself is to remain consistent with a subjective theorist's core commitments, there must also be some link between the description of the ideal circumstances for judgment and the ultimate authority of the subject's evaluative perspective. A link exists if the subject herself finds the specified conditions for good prudential judgment authoritative. The best way to establish the authority of the ideal conditions for prudential judgment is to ensure that the conditions specified are both general in nature and as evaluatively neutral as possible.[12]

The most obvious examples of sorting mechanisms subjective theorists adopt to cope with this problem are rationality and information conditions.[13] For example, many desire theorists have built

[11] Sumner, *Welfare*, pp. 160–1.

[12] This is a common description of the constraints adopted by subjective-attitudinal theorists. See, e.g. David Sobel, 'Full Information Accounts of Well-Being', *Ethics* 104 (1994), pp. 784–810, esp. p. 791; Connie Rosati, 'Persons, Perspectives, and Full Information Accounts of the Good', *Ethics* 105 (1995), pp. 296–325, esp. pp. 300–2. It is worth noting a significant difference between seeking naturalistic definitions of evaluative terms like 'well-being' and seeking a subjective-attitudinal theory of well-being. Some theorists combine the two tasks, but that is not necessary. If one seeks a naturalistic definition, then the ideal conditions must be specifiable purely descriptively, i.e. without reference to any evaluative terms. However, if one abandons this further task, then the commitment to value neutrality is simply driven by the search for widely acceptable conditions: conditions most subjects will view as authoritative.

[13] The intuition that information is what is needed to improve judgment lies behind a vast number of theories, most prominently informed desire theories of personal good.

rationality and information conditions into their theories – arguing that well-being is determined by the desires that a person would have for her not-fully-rational and not-fully-informed self, *if she were* fully rational and fully informed. Likewise, Sumner builds an information condition into his attitudinal-happiness theory of welfare. A person's assessment of her own life as satisfactory only counts as indicative of true well-being if it is informed. Of course, rationality and information are not, strictly speaking, evaluatively neutral. They are values, but they are not usually thought of as substantive welfare values. They are general epistemic values that almost everyone recognizes as improving judgment. Hence, a theory that sorts attitudes in these terms avoids specifying particular welfare values someone must embrace in order to count as doing well. At the same time, it preserves the ultimate authority of the individual's evaluative perspective because its sorting methods are ones the individual herself views as authoritative.

Let us return, then, to the question of what Celie and Quoyle lack. One thing is clear. Their problems are not straightforward problems of rationality. They are not failing to adopt the obvious means to their own short-term ends or falling prey to simple mistakes of logic. Rather, the problem lies in their perception of their situation. But is this just a way of saying that they lack full information? Would full information alone be sufficient to solve the horizon problem in these cases?

Most of the information ordinary people have, Celie and Quoyle appear to have. For example, it is highly implausible to suppose that Celie doesn't know she could leave Mr. —. She is not so isolated and lacking in common forms of knowledge that she fails to be aware of the fact that some women with abusive husbands leave or that women can, in fact, live on their own. Similarly, Quoyle is clearly aware of the fact that he could divorce Petal Bear and change the lock on the door to keep her from stealing from him.

For the purposes of theory, however, we need not limit ourselves to the kinds of knowledge most people have. Accordingly, most attitudinal theorists of well-being who appeal to information place no restrictions on the amount or type of information that is to be added. One notable exception is Richard Brandt, who in his theory of rational desire limits information to the 'propositions accepted by the science of the agent's

For example see Brian Barry, *Political Argument* (New York, 1965), chs. 10 and 11; James Griffin, *Well-Being* (Oxford, 1986), chs. 1 and 2; R. M. Hare, *Moral Thinking* (Oxford, 1981), ch. 5; John Rawls, *A Theory of Justice* (Cambridge, Mass., 1971), ch. 7; Joseph Raz, *The Morality of Freedom* (Oxford, 1986), ch. 12; Peter Railton, 'Moral Realism' and 'Facts and Values'. Richard Brandt develops an account of rational desire along structurally similar lines; and his account has been extremely influential in the literature on well-being. See Richard Brandt, *A Theory of the Good and the Right* (New York, 1998).

day, plus factual propositions justified by publicly accessible evidence (including testimony of others about themselves) and the principles of logic'.[14] This is already much more information than either Celie or Quoyle possesses. But why stop there? Of course, some critics of informed desire theories have argued that the notion of full information possessed by a single agent is incoherent.[15] If that were so, then a more restricted notion would be necessary. But for the moment, I shall simply grant the subjective-attitudinal theorist access to full information. Our question then becomes whether full factual information – however much that is – would be sufficient to guarantee that Celie and Quoyle could not fail as prudential judges.

When we consider the matter from this angle, there do seem to be significant pieces of information that Celie and Quoyle lack – bits of information that might make a difference, were they to know them. For example, while Celie may know in some sense that she could leave, she may lack any detailed knowledge of what a woman on her own would be able to do. Since (prior to meeting Sophia and Shug) she has not met independent women, she may have only a vague sense of how to manage on her own and of what troubles and difficulties would accompany such a transition. Presumably, full information would rectify this. More importantly, given Celie's low opinion of herself, she may well believe that *she* would fail at many tasks other women could handle. Presumably, however, there are facts about individuals – facts about their actual levels of intelligence, native talents, strength, endurance, educability, etc. – that entail facts about what they could do under particular circumstances. Most of us have only a vague grasp of what those facts are in our own case, and Celie no doubt vastly underestimates the facts about herself. But presumably full information would include such facts. If such facts are all Celie needs to be a good prudential judge, then information is all that is really needed to solve the horizon problem.

For reasons that will emerge shortly, I seriously doubt that the horizon problem can be fully reduced to a problem about information. Of course, information is an important aspect of the horizon problem. But it is very plausible to me that Celie might have all the facts about what she could actually accomplish, if only she tried (as well as any other relevant facts), and yet still fail to value her options appropriately. My suspicion – as yet undefended – is that the horizon problem has an evaluative dimension. If that is correct, then even huge amounts of purely factual information are not guaranteed to remove all flaws of prudential judgment. I am not unique in thinking that rationality and

[14] Brandt, *Good and the Right*, p. 13.
[15] Rosati, 'Persons, Perspectives'; Sobel, 'Full Information Accounts'.

information alone may not be enough to solve such problems. Other critics have made this point.[16] But few have specifically identified the problem as evaluative.[17]

III. EVALUATIVE PERSPECTIVES AND THE INSUFFICIENCY OF INFORMATION

I am not alone in thinking there may be an evaluative dimension to the horizon problem. Sumner also thinks that information, although important, will not suffice to solve it. A natural place to begin an investigation of this hypothesis, therefore, is with Sumner's concerns. His concerns, as well as the shape of his proposed solution, are closely tied to reflection on a particular passage in which Amartya Sen is discussing adaptive preference formation. It will be useful to have this passage before us:

> A person who has had a life of misfortune, with very little opportunities and rather little hope, may be more easily reconciled to deprivations than others reared in more fortunate and affluent circumstances. The metric of happiness may, therefore, distort the extent of deprivation in a specific and biased way. The hopeless beggar, the precarious landless labourer, the dominated housewife, the hardened unemployed, or the over-exhausted coolie may all take pleasure in small mercies, and manage to suppress intense suffering for the necessity of continuing survival, but it would be ethically deeply mistaken to attach a correspondingly small value to the loss of their well-being because of this survival strategy.[18]

On the surface at least, the people Sen describes are similar to Celie and Quoyle in that they too are unable to see just how bad the circumstances of their lives are. Somewhat ironically, however, the suggested source of the problem – suppression of distressing information – implies that a full information requirement might do the trick after all. The suggestion seems to be that these folks have the necessary resources for seeing their lives as they are, but that they have repressed this information for the sake of psychological survival. Yet Sumner thinks that Sen's cases pose an additional problem that is not reducible to a problem about information. This problem stems 'from the malleability of people's personal values'.[19]

Sumner and Sen focus on cases in which people live under oppressive conditions. Not only are their actual conditions of life hard, but many of

[16] A few examples are, David Velleman, 'Brandt's Definition of "Good"', *Philosophical Review* 97 (1988), pp. 353–71; Rosati, 'Persons, Perspectives'; Sobel, 'Full Information Accounts'.

[17] One exception is Susan Babbitt, *Impossible Dreams* (Boulder, Co., 1996).

[18] Amartya Sen, *On Ethics and Economics* (Oxford, 1987), pp. 45–6.

[19] Sumner, *Welfare*, p. 162.

these individuals have presumably been socialized in oppressive ways to accept their lot. If that is true, then they may have come to value (or at least not disvalue sufficiently) the lives they are leading.

Consider, for example, Sen's dominated housewife. She may have been raised to believe that a woman's place is in the home, answering to the authority of her husband. Suppose as well that this woman, like many women in her community, suffers from malnutrition because food supplies are limited and it is thought that boys and men must eat as they wish, with women making do with what is left. The wife is unhealthy, usually exhausted, and frequently lonely and sad. Yet she does not know much about the other kinds of lives she might be leading. Surely, it might be natural to suppose, she would prefer one of those lives, if only she had full information about what they are like.

But is this correct? Let us now suppose that she gains such information – she now knows about the kinds of lives she might lead if only she left her community. Of course, many individuals like her are, for all sorts of reasons, not free to leave. However, our concern is not so much with whether she can attain this better life, but with developing a theory that will explain the intuition that if she left, the life she could have would, in fact, be a better life. Unfortunately, these healthier, pleasanter lives are lives that do not fit with her sense of the properly subservient role of women. Given that she long ago embraced values of subservience, it is not likely that, even with full information, she will now value a life of non-subservience or agree that it would be better for her, even if she recognizes it as a possibility. Although it is plausible to think that increasing information will bring about certain changes in a person's evaluative attitudes – since a person with more information will better understand both the relations between her various ends and how to attain them – there is no reason to suppose that information alone will overturn all of her current values.

In certain cases, more information could undermine a person's commitment to particular values – for example, if it directly undermined some factual belief upon which that commitment depended entirely. If her commitment to subservience entirely depended, for example, on a false factual belief about the intellectual inferiority of women, then if this belief were overturned, she would similarly lose her commitment to subservience. But there is no reason to think that all such values entirely depend in any such direct way on factual beliefs.

Alternatively, more information might reveal a previously unnoticed conflict between the pursuit of two values, forcing a person to choose one and reject the other. For example, she might want to fulfill what she views as the proper role of a woman, but also want to live a more pleasant life. Full information might reveal to her that these wants are incompatible in her current world structured as it is and that, to

lead a pleasanter life, she would need to leave her community. But even if that were so, it is not obvious that when confronted with such a conflict, she would choose to leave. If her commitment to the properly subservient role of women is deep enough, recognizing the conflict might just lead her to reject a pleasant life. Though we may perhaps feel that a healthier, non-dominated existence must be better for her, she need not concur in that judgment given her current evaluative perspective. Thus, if one has the intuition that the dominated wife is not making a good prudential choice, one will have to acknowledge that the horizon problem cannot be solved by information alone. Any adequate solution must consider the *values* a person holds, since our values can limit our horizons as much as lack of knowledge can.

It might seem that a subjective-attitudinal theorist could not address such a problem. After all, questioning the prudential significance of values someone has embraced seems to run counter to the commitment to the authority of the subject's own perspective. However, while there is not much room for maneuver here, there is some. For if someone could be convinced that certain values had been foisted upon her in an illegitimate way or for illegitimate reasons (e.g. just so she will passively accept less in life), she might reject them for that reason alone. So perhaps the subjective-attitudinal theorist can address questions about values indirectly, by addressing questions about how a person came to have such values in the first place.

An interesting example that suggests the promise of this approach is presented in Deepa Mehta's 2005 film, *Water*, which is about the lives of Hindu widows in the 1930s. Widows, at that time, were often forced to take up a life of asceticism, poverty, and isolation – living in a special home for widows and begging for their bread. At one point in the film, Narayan, a young man intent on marrying the widow Kalyani (a taboo action), remarks to an older resident of the widows' home, Shakuntala, that these rules do not reflect some deep religious insight, but were invented simply to allow families to be spared the expense of feeding and caring for widows. Although Shakuntala is at first shocked by this suggestion, its effects on her and her self-understanding of her own values are profound. By the end of the film she is willing to ask Narayan to help Chuyia, a seven-year-old child widow, to escape from the home to a better life.

Stories like this suggest that what is significant about Sen's dominated wife is precisely the fact that she is *dominated*. The problem has to do with how and why she has come to hold the values she does. If she herself understood the how and why, perhaps she would think differently. Sumner's theory follows up on this line of thought. I shall now examine his approach and consider whether it can solve the horizon problem.

IV. SUMNER'S SOLUTION: AUTONOMY-AS-AUTHENTICITY

According to Sumner, well-being consists in 'authentic happiness'. 'Authentic' signals that not all self-ascriptions of happiness or satisfaction count as genuine indicators of well-being. To count, such ascriptions must satisfy two further conditions. The first is an information condition, as mentioned previously. The second is that whatever values inform the individual's assessment must be autonomously held.

In its root sense, 'autonomy' means self-rule. But a number of distinct conceptual strands are associated with the term. Sumner wishes to invoke the authenticity strand – the idea of acting from values that are truly one's own (in some yet-to-be-specified sense). His thought is that authenticity will rule out cases in which a person's current values are alien to the person herself in an important sense.

Sumner is also drawn to authenticity because he hopes it will provide a sufficiently neutral, and hence sufficiently authoritative, way of sorting evaluative attitudes. On the surface at least, an autonomy condition – in contrast with rationality and information conditions – poses more potential for controversy. 'Autonomy' is the name for a complex set of (non-epistemic) values that many non-liberals, and many people from non-Western cultures, view with suspicion. However, in fairness to Sumner, the meanings of the word that are most widely viewed as problematic – such as the idea of living in a highly independent way that challenges common societal norms – are not connected in any strong way with the meaning to which he appeals. It is possible to hold even the values of subservience and obedience authentically. Hence, if any strand in the notion of autonomy is likely to have wide appeal, the notion of authenticity – with its implication that the values one embraces are truly one's own – is plausibly it. Finally, as Sumner recognizes, much will turn on how the authenticity condition is described. If a theory of authenticity can appeal to the judgments of the individuals in question – to their own sense of which values are alien to them and which are not – then it may succeed in maintaining the appropriate kind of link with the subject's evaluative perspective.

Sumner does not officially commit himself to any particular theory of autonomy-as-authenticity. Rather, he insists that what he wants is a new hybrid theory that incorporates the best aspects of several distinct strands of theorizing. What matters for our purposes is simply that (1) Sumner wants to incorporate some form of authenticity requirement and (2) he thinks the most promising approach is John Christman's.[20]

[20] Christman, 'Personal History'.

Importantly, he wants a theory that, like Christman's, is procedural, i.e. one that emphasizes the origin and development of values. For this reason let us briefly consider Christman's view.

Christman begins from the plausible enough thought that 'what is crucial in the determination of the autonomy of a desire [value or other action-shaping attitude] is the manner in which the desire was formed'.[21] Since not all ways of acquiring values are bad, and since 'all of our aims, values, and ideals will have been influenced to some extent by our peculiar histories and the socialization processes that have shaped us',[22] we need a way of distinguishing the good processes from the bad. To do this, Christman introduces a counterfactual test for authentic values: For any agent A who values X, we must consider the development of A's valuing of X and ask several questions. First, we must know whether X was adopted under conditions that make it impossible to reflect on what is happening. If not, then we must ask whether A would have resisted adopting X, had A been attending to the process as it occurred. If the process *was* a reflection inhibiting one, then we must ask the more complex counterfactual question whether A would have resisted the acquisition of X, if A had not been in reflection-subverting conditions *and* had actually attended to the process. Only if A would not have resisted coming to hold X under either set of counterfactual conditions can we view A's choice to embrace value X as an authentic expression of A's preferences.

In more recent work, Christman has emphasized that what matters is the person's attitude towards a particular value given its etiology, as opposed to an attitude to its etiology alone. It is possible to believe that one acquired a good value through bad means. He also now emphasizes, in place of talk of simply rejecting a value, talk of experiencing an 'affective sense' of alienation from the value in question. However, neither of these amendments matters much for our purposes, as the theory is still of the same broad type.[23]

Our concern is with whether Christman's theory or one suitably like it can help us build a better theory of welfare or well-being. Yet, promising as Christman's device may seem, it faces problems that matter significantly for our project. His theory assumes that, under normal conditions, individuals are able to recognize 'bad' processes of socialization when they see them. It appeals to the agent's own (actual or counterfactual) judgment about such processes, allowing

[21] Christman, 'Personal History', p. 10.
[22] Sumner, *Welfare*, p. 170.
[23] John Christman, 'Autonomy, Self-Knowledge, and Liberal Legitimacy', *Autonomy and the Challenge to Liberalism*, ed. J. Christman and J. Anderson (New York, 2005), p. 334.

that judgment to be authoritative. However, what a person sees as a process worthy of resistance will depend in part on her current values. Admittedly, the whole point of Christman's theory is to focus on how people came to acquire certain values in the first place. Thus, he asks us to consider the developmental history of particular desires or values. But presumably, even if we consider some point in a person's history before a particular value has been adopted, this does not mean that no values inform the person's outlook at that earlier time. Hence, the desire whose origins interest us will not have been formed in an evaluative vacuum.

To understand my concern, consider again Sen's example of the dominated wife. Let's assume that she began life as the daughter of loving parents in a culture with deep traditions of female subservience. If this sounds paradoxical, it is worth remembering that love and oppression are not automatically opposed. Moreover, love, as we shall see, is importantly related to trust. Let us also stipulate that she was brought up to accept subservience in a way that we would regard as oppressive. For example, although a minority group living in the same geographical area does *not* hold similar views about the properly subservient role of women, she was forbidden from the time she was a small child to have anything to do with them. She was not to talk to them or to play with their children. Moreover, from an early age, she was sent to a special school where the lessons of female subservience were taught on a daily basis and where young women who showed signs of independent thinking were severely punished. Both Christman and Sumner would classify such conditions as oppressive and would thus have concerns about the autonomy of any values so formed.[24] Despite all of this, it is still plausible to imagine that, at the time, the young woman did not resist her education. More significantly, it is plausible to imagine that she would not have resisted even if she had stopped to reflect upon it as it was happening and even if this reflection took place under conditions not limited in the way her current circumstances are.

Of course, since we are trying to understand the perspective she had prior to adopting beliefs about the inferior status of women, it is important that we do not explicitly or implicitly attribute such a failure of resistance to these kinds of beliefs. But we do not have to. We need only imagine that she trusts those around her and that she lacks certain other values that we have and which Christman's theory tacitly presupposes. His theory assumes that, prior to developing 'bad' values through oppressive socialization, resistance to such values will be automatic because there will be nothing there to stop it. But one

[24] Sumner, *Welfare*, p. 170; John Christman, 'Liberalism, Autonomy, and Self-Transformation', *Social Theory and Practice* 27 (2001), pp. 185–207, at p. 203.

could just as easily point out that there will be nothing there to support it either. Resistance to practices like the ones I described above only seems natural to us because we are already advocates of open forms of inquiry. We already hold values that dictate that it makes sense to resist any attempt to teach us something when the teaching is entirely one-sided and that it makes sense to resist any attempt to limit our access to other people with different views.

Yet there is no reason to assume that our young woman already possesses such values. Moreover, if she trusts those around her, then she may be vulnerable to their tactics in a way that Christman's device simply cannot detect. Given this trust in her parents and teachers, she will not begin with the assumption that the education she is receiving – despite its intensity and the harshness of the punishments meted out to some girls – is something one ought to resist. Nor will she start out with the assumption that being kept from talking with a certain group of people is a bad thing – an oppressive maneuver designed to keep her from learning truths. Finally, all we need is for this to be a description of a possible case. (Indeed, my example of Shakuntala was intended to acknowledge that there probably are some cases that work the way Christman intends.) If it is a possible case, then Christman's theory of authenticity will fail to eliminate all of the values we intuitively view as inauthentic; and it will fail, in turn, to help Sumner eliminate all cases of bad prudential judgment that flow from inauthentic values.[25]

It may be tempting to suppose that my argument has neglected the way in which Christman's device is supposed to work when included as a part of Sumner's larger theory of welfare. Since it is Sumner's theory that we are really examining, this would be a problem if it were true. For example, it might seem to be relevant that in Sumner's theory the autonomy requirement works in conjunction with a condition of full information. But even so, Sumner's information condition will not necessarily help to eliminate problems that Christman's device alone leaves untouched. Consider again the woman I described earlier, once she is grown up and living, lonely and sad, as the dominated wife who must eat less than she needs for health. The theoretical question is whether this same woman – exactly as she is now, except possessed of full information and without the influence on her judgment of any

[25] I find it plausible to suppose that values formed under such conditions are inauthentic. It is possible that Christman might insist that if she did not feel alienated from these values, even after reflecting on their origin, then the values are authentic. However, it is not important for me to take a stand on this issue here, since what primarily interests me is the role Christman's theory might play in a theory of welfare. Whether or not we think her values are authentic, it is plausible to assume that the woman described is not making prudentially good choices. And that is sufficient to establish that Christman's theory will not help Sumner.

values she currently possesses that fail Christman's test – would be able to offer a more realistic appraisal of her life. As we have seen, the values she acquired as a child will not be eliminated because they pass Christman's test. So they can still influence her judgment. As concluded before, now that she has these values and has lived with them for a while, it is unlikely that even full information about how other women live will lead her to reassess the relative goodness of her life in a way that makes sense to us. Hence, the problems for Christman's theory really are problems for Sumner's project as well.

As I have described her, the dominated wife poses a problem for the theorist of welfare who adopts an authenticity condition. However, the example does not directly challenge the relevance of authenticity to a theory of welfare. What it does is show that, as yet, we have no reliably authoritative way of distinguishing between authentic values and inauthentic ones. The dominated wife's values presumably are inauthentic. The problem is that Christman's device in Sumner's theory fails to deliver this intuitive verdict. Moreover, the prospect of finding some other neutral and equally authoritative device that can properly distinguish authentic from inauthentic values looks dim. What gave Christman's approach a chance at being authoritative (in the right sense for Sumner's theoretical purposes) was precisely its reliance on the individual's own perception of her own socialization. But, as it turns out, this same feature is the root of the problem. Given how subtle oppressive forms of socialization can be, it will be difficult to describe another method that catches all and only the inauthentic values in a way that subjects themselves can accept as authoritative.

The current example raises concerns about our ability to fashion a requirement that will detect inauthenticity. However, I want now to turn to a deeper worry about Sumner's theory and, by extension, all subjective-attitudinal theories of welfare. I want now to suggest that there are good reasons for thinking that even information plus authenticity will not be sufficient to solve the evaluative aspect of the horizon problem. For, as I will now try to demonstrate, these two conditions alone do not help us to solve the kinds of problems posed by individuals like Celie and Quoyle.

V. A SENSE OF SELF-WORTH

One particularly striking feature of both Celie and Quoyle is their low sense of self-worth. As I suggested before, after years of hearing Mr. — put her down, Celie *feels* worthless. Without a proper sense that she, Celie, matters, her idealized self will be unlikely to give her non-idealized self good advice. Nor will she be able to assess her life fairly in comparison with other possible lives. Although she may recognize

that other lives are happier or more successful, she will not necessarily see them as better for her if she herself feels unworthy of the lives in question.

Admittedly, the phrases 'self-esteem' and 'sense of self-worth' are used in a frustratingly wide range of ways. Despite appearing to refer to a single psychological attitude, these terms are probably best understood as referring to a cluster of semi-related psychological attitudes. Moreover, some of the items in the cluster behave more like straightforward evaluative beliefs, while others are affective, and so seem to be at least partly independent of the forces that operate on ordinary beliefs. Still, many people recognize the central elements of this cluster and recognize the crucial way in which they contribute to well-being. My suggestion is that not only do these elements contribute to well-being directly, by making one feel better, but they also contribute to good prudential judgment. So unless a subjective theory has a way of correcting for their absence, there is no guarantee that the judgments of the idealized self will be a good guide to the individual's well-being.

It might be tempting to think, again, that the full information condition would solve this problem, albeit indirectly. After all, Celie has a low sense of self-worth because she also believes that she is unlovable, clumsy, stupid, and lacking all talents. Full information would reveal to her that she is not stupid at all, that many people are quite capable of loving her, and that she is a talented seamstress with a fine fashion sense as well. Presumably learning this type of information would boost her self-esteem. However, while that might happen under certain circumstances, it will not always happen. A low sense of self-worth is not generally dependent on other factual beliefs about oneself. Even in cases like Celie's, in which false factual beliefs help to explain the origins of low self-esteem, it is well known that correcting the beliefs does not always lead to a change in the feelings one has about oneself. The sense of self-worth is (at least partly) affective and so potentially immune from normal mechanisms of belief revision. This is what explains why highly talented people with low self-esteem can spend years in therapy trying to recover it. Despite many actual successes in their lives that overturn any false beliefs they may have about their talents, they continue to feel badly about themselves and this may continue to influence their judgment.

Low self-esteem is sometimes also the product of psychiatric conditions such as depression. So it might seem that some kind of condition ruling out psychiatric conditions in the idealized self would do the trick. But while those who are depressed typically have low self-esteem (they typically evaluate everything in a negative light, the self being no exception), low self-esteem is not limited to cases of clinical

depression. It occurs in many people with no recognizable psychiatric illness at all.

What, then, about authenticity? Could we try to establish that the values adopted by people with low self-worth are inauthentic? Is there a useful parallel between Celie and Sen's dominated wife? To begin with, Celie's sense of worthlessness is not the direct result of oppressive socialization into general values of subservience. She probably holds no views about the inferior role of women or the goodness of subservience generally. So, unlike the dominated wife, who submits to ill-treatment out of a sense that her husband has a right to treat her that way, Celie submits, at least initially, because she doesn't know how to escape and doesn't know whether anything better is possible. Later, we may assume, she remains because she has come to see herself as worthless and this shapes her assessment of the few options she does know of.

Still, one might hope that a device like Christman's could be used by a theorist of welfare to cover cases of violent abuse like hers. Christman's idea seems to be that whenever knowledge of the origin of an attitude would, by itself, be sufficient to make an individual (at least try to) resist the development of the attitude, then the attitude in question is inauthentic. Perhaps if Celie, as she was *before* she developed low self-esteem, could be brought to understand that her later self only feels so worthless because of all the abuse, verbal and physical, heaped on her by Mr. —, then she would try to resist developing low self-esteem. If that were true, then perhaps it would make sense to say that her low self-esteem is inauthentic. However, it is not *obvious* that she would resist, even if she knew this. As with the dominated wife, that will depend on the evaluative perspective she had prior to developing low self-esteem and on the kinds of resources that perspective contained.

Much more importantly for our purposes, however, is the fact that there are numerous cases of low self-esteem in which it makes no sense to view the attitude in question as inauthentic. So even if full information plus authenticity will handle Celie, it will not handle all cases of low self-esteem. To see this, consider the interesting psychological trait psychiatrists label 'rejection-sensitivity'.[26] This is presumably a biological trait: something people are born with. People who have it experience greater discomfort in the face of negative feedback – they feel worse than normal people when someone is rude to them, gets angry with them, or simply ignores them when they are seeking attention. On experiencing any kind of loss, 'these people feel more pain or come closer to depression than do most men and women'.[27] In addition, because they are vulnerable to emotional pain in ways that

[26] Peter Kramer, *Listening to Prozac* (New York, 1993), ch. 4.
[27] Kramer, *Listening*, p. 71.

others are not, rejection-sensitive people grow to fear certain kinds of interactions. They may (unconsciously) adopt various behaviors as part of an overall strategy of self-protection and pain avoidance – behaviors that seem eccentric or excessive from an outsider's perspective. It is important to emphasize that what I am describing is a trait people simply have or lack. Moreover, it results in a psychological profile that is only a minor variant on normal functioning, so we cannot say this is a trait that leads to mental illness. People who have it are fully functional members of society. Indeed, the trait is relatively common, and comes in varying degrees ranging from strong to mild forms. Yet possession of this trait sets people up in various ways for developing low self-esteem.

To see what I have in mind, let us retell Quoyle's story somewhat. Let us leave the explicit abuse out of his past, but imagine instead that Quoyle is highly rejection-sensitive and that in his family it is harder than usual to gain positive feedback. What encouragement is on offer is tied to certain values and attitudes about masculinity and appearance. In short, the parents favor individuals who are good-looking and assertive 'go-getters'. Unfortunately for Quoyle, he is awkward, unattractive, and naturally shy. In an effort to fit in and avoid the painful experiences of disapproval, Quoyle tries to attend to his appearance and to be more assertive. In the process, he internalizes his parents' flawed attitudes. However, because he has little control over his awkwardness and shyness he is doomed to receive little in the way of positive feedback from his parents. Somewhat ironically, a deep need to avoid the pain of rejection leads Quoyle to internalize a set of attitudes that set him up for more pain later on. We can imagine that Quoyle is now caught in a destructive psychological cycle. The very attitudes in terms of which he assesses himself – those upon which his esteem depends – are ones in terms of which he cannot succeed. Eventually, if his self-esteem is completely eroded, he may end up adopting a fatalistic outlook on his prospects in life. In short, he may end up in the position of the original Quoyle – an easy target for a predator like Petal Bear.

It is important to be clear about how the case works. The attitudes Quoyle adopts about what a man must be like in order to be successful, worthy of attention, love, respect, etc. are one source of the problem. However, the deeper explanation of why he adopts these is that he is unusually vulnerable to the opinion of others who hold them. This leads, in turn, to a cycle of failure, pain, and ever deepening self-loathing that now has a life of its own. It will not help now to learn that there are people in the world who care for men who are not good-looking 'go-getters'. This example provides an illustration of the cases usually overlooked in discussions of the horizon problem. It is similar to some of our earlier cases insofar as Quoyle's defects in judgment can be traced

to his evaluative perspective. But it is different because the elements of this perspective cannot be traced to oppressive socialization processes or even explicit abuse. Another individual in the same circumstances who lacked this trait would not have had his self-esteem eroded.

To understand the problem here, one must take notice of features of the agent as well as features of the environment. But this suggests that the resultant attitude is authentic, since it partly derives from who Quoyle is. Of course, authenticity is notoriously hard to define, because it is difficult to give an account that strikes a plausible balance between being too demanding and being too lax. It is tempting to note that Quoyle may have been born with a trait, but that he was not born with low self-esteem. Yet, as Sumner himself emphasizes, we cannot demand that all adopted or acquired attitudes be counted as 'alien', not even all attitudes adopted or acquired from others.[28] After all, all of our evaluative attitudes are acquired in some sense.

The most obvious way to try to define authenticity, then, is to focus on suspect processes of acquiring values or attitudes. But if we grant that this is the most promising approach to authenticity, then we will have to admit that there is nothing abnormal or morally suspect about the process that leads Quoyle to develop low self-esteem. The attitude and its adoption are intimately connected to a trait he naturally possesses. Even though his acquiring low self-esteem was a contingent matter, the fact that Quoyle's own nature partly explains the ease with which he falls into it should make us hesitate to say it is inauthentic. It is also important to note that one need not accept the particular details of this psychological example. One need only accept that some such psychological features of human beings work in roughly the way I have described.

If full information, authenticity, and lack of mental illness are not jointly sufficient to block low self-esteem, and if low self-esteem is a threat to good prudential judgment, then the subjective-attitudinal theorist faces a serious problem. As there is apparently no indirect way to eliminate low self-esteem, it seems that a reasonable level of self-esteem will have to be built directly into the account of the ideal conditions of judgment. But this seems to violate the subjective-attitudinal theorist's commitment to the authority of the subject's own perspective.

Perhaps, it might be argued, enough people recognize the value of healthy self-esteem that adding a self-esteem requirement to the theory will not violate the requirement that the ideal conditions of prudential judgment be acceptable to the subject herself. Perhaps, for most

[28] Sumner, *Welfare*, p. 170.

subjects, such a condition would be accepted as authoritative. However, I doubt that anywhere near as many people accept the importance of self-esteem for prudential judgment as accept the importance of information and rationality. Moreover, we will presumably need not just to stipulate that a person has a sense of self-esteem, but we must stipulate that she has an appropriate *level* of self-esteem. But what level is appropriate? How much must one value oneself to prevent lack of self-esteem from threatening one's prudential judgment? Our intuitions about this are largely based on our observations of the choices people with differing levels of self-esteem make. If we view someone's choices as prudentially bad, then we may well regard their level of self-esteem as problematic. But this suggests that there will be no way to agree on the level of self-esteem necessary for good judgment without first agreeing on many other substantive evaluative matters. Hence, even if almost everyone can agree that extremely low self-esteem is prudentially bad, problems remain for the subjective theorist of well-being committed to the authority of the evaluative perspective of the individual.

VI. DIM PROSPECTS FOR SUBJECTIVE-ATTITUDINAL THEORIES AS CURRENTLY CONCEIVED

I have argued that low self-worth can contribute to poor prudential judgment; and that there is no obvious way to eliminate such evaluative attitudes from ideal prudential judgments while honoring the subjective-attitudinal theorist's distinctive commitments. Authenticity, even if important, is not sufficient to solve the horizon problem because, in certain cases, authentic values can be as problematic as inauthentic ones.

Indeed, with this point now in front of us, we can see that a similar point could be made using the earlier example of the dominated wife. As I originally told her story, her parents, although loving, nonetheless employed oppressive means of educating their daughter. However, we could retell the story yet again, and imagine that they did not employ oppressive means, but that because female subservience was something they valued and she trusted them, she still came to value a life of subservience. This is now not a story about oppressive socialization, but simply a story about ordinary, everyday socialization – albeit into values that turn out to have serious consequences for the woman's ability to live well later on. Like low self-worth, which can be authentic yet problematic, so too the dominated wife's values may be authentic but problematic.

It is nonetheless important for my purposes that this point about the insufficiency of an authenticity condition was initially made with the

example of low self-esteem. We need to avoid the temptation to look for alternative accounts of authenticity – accounts that count the wife as I just now described her, the wife who did not grow up with obvious forms of oppressive socialization, as having inauthentic values. Part of what I wish to emphasize is that a more complex notion of authenticity that counts more forms of socialization as oppressive would not really be a helpful move in this context. The exploration of the horizon problem to date has been too much focused on the social or cultural conditions that lead people to adopt poor prudential values. Not enough attention has been paid to the deeper psychological roots of the problem. Reflection on these deeper roots, however, makes a solution that meets the standard constraints of subjective-attitudinal theories even less likely.

Might it be objected that the intuitions appealed to here are intuitions a subjective-attitudinal theorist cannot take seriously? If the values the dominated wife adopts seem problematic even when they are authentic and informed, then perhaps only an objective commitment to the general badness of subservience could be driving the intuition. Yet (it might be argued) those are precisely the kinds of intuitions by which subjective-attitudinal theorists think we should not be moved. From the subjectivist point of view, no value can be ruled out *a priori* as incompatible with a good life. Whether subservience is compatible with a good life is a fact that must emerge from facts about the psychology of particular individuals. Whether or not particular values are prudentially flawed must depend on facts about the agent's own capacities of appreciation.

I believe this worry is ungrounded. But to see why, it is useful to distinguish between two ways of thinking about the cases in question. Some theorists do have the anti-subjectivist intuition that an item, X, could count as good for a person, despite her being completely unable to appreciate X's goodness. They hold that the presence of X in a life makes a life better period, independently of whether the agent whose life it is sees it as better. Such intuitions are indeed antithetical to the whole subjectivist approach.

However, a different intuition is that the dominated wife herself would appreciate the pleasanter, healthier life if she were living it and held the values appropriate to it. In other words, the life would have evaluative authority for her from within. Described this way, the intuition is not divorced from the thought that what counts as good for her depends on her particular capacities of appreciation. The problem is that one can have this intuition and yet not be able to find a way to defend the appropriate perspective – the internal perspective of that alternative life – to the wife as she is now. The life we intuitively select as better for her is one that would have a kind of internal authority for her if she were living it, but that has no external authority she can

recognize from her current position. Moreover, what seems to drive the intuition in a case like this is not simply the fact that we have before us a person whose values we find odd, but that we have before us a person whose values make her incapable of appreciating the disvalue in acute suffering. Such an intuition strikes me as, fundamentally, subjectivist in that it focuses both on the individual's capacities for appreciation and on the individual's subjective experience. However, subjectivist or not, such intuitions cannot, apparently, be captured by theories that demand that the authoritative attitudes are those chosen in a manner the individual herself could recognize as an improvement. Certain people inhabit evaluative perspectives that make it impossible for them to appreciate from their current position what a good life might be.

VII. CONCLUSION

In this article, I hope to have persuaded readers of several significant claims. First, there is an *evaluative* dimension to the horizon problem. Second, while we may care that our values be authentic for other reasons, adopting authenticity as an additional filtering mechanism will not help. Authentic values can be just as problematic as inauthentic values. Moreover – and this is the most important point – some of the problematic cases have incredibly deep psychological roots, so that no mere focus on socialization processes will help.

As a result, the prospects for solving the horizon problem in terms that satisfy the subjective-attitudinal theorist's constraints are dim. As long as subjectivists hold on to the requirement that the individual's current perspective be authoritative – even if only indirectly in terms of specifying which attitudinal filters may be employed by a theory – the horizon problem will remain.

My own intuitions suggest that, rather than abandon the subjectivist project, subjective-attitudinal theorists should reconsider the adequacy of the constraints that have guided their theories thus far. In particular, they should reconsider the kind of authority that is granted to the individual's own current evaluative perspective. Some theorists may insist that the focus on individual 'sovereignty' is definitive of subjectivism. However, as I see it, a more natural reading of the subjectivist commitment is that it is a commitment to building a theory entirely out of individualized psychological materials (no appeals to purely objective values) combined with a commitment to the thought that an acceptable theory cannot classify a life as good, if living that life is not internally good for the person living it. In other words, it is a commitment to the thought that a good life, or at least certain key

features of a good life, must be appreciable as such from within.[29] But we needn't accept the opposite idea: that a bad life must be appreciable as such from within. Since, as I said before, certain people inhabit evaluative perspectives that make it impossible for them to appreciate from their current position what a good life might be, we should not limit ourselves to definitions of their good that they will appreciate now.[30]

hawkins@utsc.utoronto.ca

[29] I do not mean that subjectivists should reinstate an experience requirement. Some things that add to the goodness of a life might be outside awareness. All I mean is that a subjectivist would reject a theory that says a life is good for a person even though the person who is currently living that supposedly good life neither feels good about it nor judges it to be going well.

[30] For helpful feedback on earlier versions of this article I would like to thank members of the 2004 Toronto Workshop on Well-Being and Autonomy and members of the 2004–5 Faculty Fellows Seminar at Harvard University's Edmund J. Safra Center for Ethics. I would also like to thank the following individuals for particularly helpful comments and encouragement: Wayne Sumner, Connie Rosati, Jodi Halpern, Simon Keller, and Norman Daniels.