Rational Choice, Changes in Values over Time, and Well-Being

DENNIS MCKERLIE
University of Calgary

Sometimes we make decisions which affect our lives at times when we will hold values that are different from our values at the time the decision is made. What is the reasonable way to make such a choice? Some think we should accept a requirement of temporal neutrality and take both sets of values into account, others think we should decide on the strength of our present values, yet others think that in evaluating what will happen at that other time we should use the values that we will endorse at that time instead of our present values. These views see the problem as one about finding some attitude towards time itself that is distinctively rational. This *article* argues that these views are subject to serious objections. It suggests that instead we should think in terms of well-being. If a person approves of, or positively responds to, the way their life is going they will experience more well-being than if there is no positive response. The *article* explores the implications of a positive response condition on well-being for deciding what it is rational to do in cases involving changing goals.

I

Typically our goals and values change during our lifetimes. What I think important and try to achieve during one stage of my life might differ radically from what I care about most at some other time. The change can happen in different ways and for many reasons, but a familiar case is that people often have different aims when they are middle-aged and when they are elderly.

These changes can create conflicts. Suppose that my present goals and the goals that I will endorse when I am old are significantly different. Choices that I could make now might help me to realize my present goals but hinder my ability to eventually realize my future goals. This might be so because the resources available for me to use over my lifetime are limited. However, in speaking of conflict I will assume not just that there is competition for scarce resources but that there is a direct disagreement about value between myself now and myself at the later time. I now believe that goal G1 is more important than goal G2, but at that future time I will instead think that G2 is more important than G1.[1] Although I will consider other examples, the

[1] In his discussion Nagel emphasizes that he is concerned with changes of values: 'It may happen that a person believes at one time that he will at some future time accept general evaluative principles – principles about what things constitute reasons for

© 2007 Cambridge University Press
doi:10.1017/S0953820806002342
Utilitas Vol. 19, No. 1, March 2007
Printed in the United Kingdom

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

basic case I will discuss is a conflict between what my current goals would mandate if they were applied to my life at the future time T2 and what the goals I will hold at T2 itself would recommend about my life at T2. I call this a 'present–future' conflict.[2]

The conflict raises an issue about how we should think when we make a present decision that will affect our future. The issue is thought to be one about the nature of prudence, on the understanding that prudence is precisely the kind of thinking that rationally considers the impact of choices on one's life as a whole. How should prudential thinking respond to a change in goals and values? Should I simply use my current goals for a decision that is to be made now, or just my future goals when it is a matter of evaluating the facts that will obtain at that future time, or is it more reasonable to somehow take both sets of goals into account?

My discussion will be governed by an assumption that I will call 'modest objectivism'. I will assume that some goals can be more important or more valuable in themselves than others, and that people can make mistakes when they make judgments about the importance of specific goals. I think this assumption is reasonable, but I will not argue for it here. However, it does distinguish my view from views that refuse to question or challenge the preferences that people have at particular times.

I call the assumption 'modest' objectivism because I do not suppose that appealing to the objective value of goals is enough to solve the problem I am discussing. That is, we should not simply say that the reasonable choice for the person in the present–future case is to further the goal that has the greatest objective value, regardless of whether it is a goal that she now accepts or a goal that she will eventually accept in the future. Later I will explain why I do not think this response is sufficient.

To begin with I will explain, and criticize, three responses to the problem of conflicting goals. These views share a certain assumption about the nature of the problem. They suppose that the basic issue is about time and the intrinsic importance of time for rational decision-making. Consequently, they suppose that a solution will identify some attitude towards the significance of time as reasonable. Once this basic

action – which he now finds pernicious. Moreover he may believe that in the future he will find his present values pernicious' (Thomas Nagel, *The Possibility of Altruism* (Oxford, 1970), p. 74). When I speak of goals I mean goals which depend on evaluative beliefs. Mere preferences independent of evaluation are excluded, although they provide goals in the ordinary sense. I will use 'goals' and 'values' virtually interchangeably.

[2] I will assume that the agent thinks it is appropriate to apply his present goals to a future time at which he no longer endorses them. In Parfit's terminology, his allegiance to his present goals is not 'conditional on their own persistence' (Derek Parfit, *Reasons and Persons* (Oxford, 1984), p. 151). He believes at T1 that his life at T2 should match those goals even though he realizes that at T2 he will no longer want his life to be that way.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

attitude has been identified, it will tell us whether we should take into account all of our goals, or only our present goals, or only our predicted future goals when it is a matter of evaluating facts or events at that future time.

I will argue that none of the three solutions is convincing (sections II–IV). They face intuitive objections and more theoretical objections. I will suggest that this should motivate a different approach to the problem. Instead of attempting to identify some attitude towards time itself as uniquely rational, we should regard the problem as being (at least partly) a matter of understanding the nature of well-being (sections V and VI).

I will argue that a particular view about well-being – a view which I will call the 'positive response condition' – provides the best explanation of our judgments about which choices are reasonable in cases of changing goals. If we accept this condition a particular choice will be reasonable, not because it correctly responds to the rational significance of time itself, but because it is the choice that will maximize the agent's well-being.

II

The first view might seem to be the solution offered by traditional theories of prudence, although not all defenders of the rationality of prudence see it as having this implication about changes in values. Because of this connection I will call it the 'prudential solution'. The basic claim is that *decision-making*, in order to be rational, should be governed by a strong requirement of temporal neutrality. Mere differences in timing, or facts about when things occur, cannot in themselves affect our reasons for acting, so we should ignore those facts when making a decision.

This idea is thought to lead to the conclusion that we should take all of our goals equally into account when making a decision that will affect our lifetime, regardless of the particular times at which we might endorse those goals during our lives. In my example it would be a mistake for the person at T1 to take into account only her goals at T1 about the later time T2, but it would be equally a mistake for her to decide at T1 merely on the basis of the goals that she will eventually believe in at T2. Both those decisions would involve an objectionable bias in favour of some particular time. Temporal neutrality requires that she should be impartial between her current goals for her life at T2 and the goals that she will have when T2 arrives. She needs to strike some kind of balance between the conflicting goals.[3]

[3] Parfit apparently sees the traditional way of understanding prudence and the requirement of temporal neutrality as leading to this conclusion. 'On the Self-interest

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

One concern with this solution is that it clashes with modest objectivism. If some goals can indeed be more important than others, it is difficult to understand how there could be a requirement for an agent to give equal weight to all of the different goals that she might espouse during her lifetime. After all, her goals at T2 might actually be much more valuable than the goals that she held for T2 earlier in her life at T1, even though she herself did not recognize this at T1. In that case it seems that it would not be reasonable for her to give her goals at T1 the same weight as her goals at T2.

Of course this objection will not worry someone who rejects objectivism. And it might be suggested that this strong form of temporal neutrality would only be endorsed by someone who did not believe that goals could differ in objective value. Since I find modest objectivism persuasive I do not see this as a satisfying reply to the objection.

A more complicated reply to the objection would distinguish between what is objectively reasonable and what is subjectively reasonable. Modest objectivism about goals might be part of an answer to the first question. If goal G1 is objectively more valuable than goal G2 that is a reason for thinking that it is objectively more rational to choose to realize G1 than to choose to realize G2. But we can also talk about what it would be subjectively rational for an agent to choose. Subjective rationality will depend on the agent's beliefs about what reasons she has for acting, whether those beliefs are true or false. This reply to the objection suggests that the prudential solution is put forward as an account of what it is subjectively rational for an agent to choose. It is unfair to criticize the prudential solution by appealing to modest objectivism about goals, because modest objectivism is relevant to what it is objectively rational to choose.

However, the prudential solution does not seem to be intended merely as a theory about what is subjectively rational. And if it were it would not be a plausible theory. If subjective rationality is understood in terms of a person's beliefs about what it is rational to do, those beliefs might mandate that she should use her goals at T1 to make decisions about T2, rather *than* giving equal weight to both sets of goals. If she believes at T1 that it is rational to decide using the goals she holds at T1, how

Theory, this young man [Parfit is talking about the agent in an example of Nagel's] must give the same weight to his present and his predicted future values and ideals' (Parfit, *Reasons*, p. 155). The Self-interest Theory goes beyond temporal neutrality because it claims that the ultimate object of rational concern is our own interest. But Parfit seems to think that it is the Self-interest Theory's acceptance of temporal neutrality that commits it to this conclusion.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

can it be subjectively rational for her to give equal weight to both sets of goals?[4]

Modest objectivism is not the main reason why people have felt dissatisfaction with the prudential solution. Thomas Nagel, who passionately defends the rationality of prudence in general, is not completely satisfied with the prudential solution.[5] He comments that one might respond to changes in goals over time in this way, but that one also might respond by relying on one's present goals and values to make decisions about the future. He points out that if prudence adopts the solution I have described it could be said to treat goals and values as if they were merely preferences by assigning them equal weight. Nagel apparently does not see this point as a criticism of the solution, but Derek Parfit does regard it as a damaging criticism.[6] Since accepting something as a goal involves a commitment to its value or importance, it might be unreasonable to find a compromise by assigning equal weight to values that you endorse and values that you (currently) reject.

Also, the prudential solution's conclusion about present–future examples seems intuitively implausible. In Nagel's example a person who now values spontaneity, risk, and strong emotion knows that twenty years in the future he will value security and tranquillity. Suppose that he is considering a choice that he can make now that would compel *him* to embark on a dangerous adventure twenty years from now. In making his decision it can seem at the very least odd that he should be rationally required to give equal weight to his current values, given that he will have abandoned those values when that future time arrives.

However, the intuitive case against the prudential solution has usually been made with a different example. Parfit thinks it is clear that he should now give no weight at all to the view he held about his future when he was an adolescent, the view that he should spend his life as a poet.[7] This is a case of past–present conflict. Some people will find the example so compelling that they would think it is a condition

[4] It might be denied that what it is subjectively rational for an agent to do at a particular time must be explained in terms of the beliefs about reasons for acting that the agent holds at that particular time. Perhaps it can also depend on goals and beliefs about reasons for acting that she holds at other times. In that case subjective rationality might require her to give weight to her future goals as well as her present goals, even if she herself does not now believe that this is what it is rational for her to do. I will not argue against this suggestion, but I think it will be difficult to make a convincing case that the prudential solution is the best account of subjective rationality.

[5] Nagel, *Possibility*, p. 74 fn.

[6] Parfit, *Reasons*, pp. 154–5.

[7] For the purposes of his argument Parfit supposes that this example does not involve a change of values (*Reasons*, p. 157). But if we suppose that there is a change of values the example retains its force.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

on a successful response to the problem of changing goals that it should deliver this conclusion. The prudential solution fails this test.

Finally, it is not clear that the prudential solution is supported by temporal neutrality. That idea means, roughly, having an equal concern for every temporal stage in one's life. But the person who uses her current goals to make decisions about a later time when her goals will be different is not necessarily violating that requirement. It is not that she doesn't care about her distant future, or cares about it less than she cares about the present. She cares about her future a great deal. That is why she thinks it is important to use the right set of goals in making decisions that will profoundly affect her life at that time. Her mistake, if we think that it is a mistake, is in letting her present goals represent her concern about a future time when she will be committed to different goals. And that mistake does not obviously involve a pure time preference.

It might be claimed that it is a violation of temporal neutrality at a higher level. She chooses to follow her present goals rather than her future goals, and that is just a more complicated form of a bias in favour of the present. But it is questionable whether this is the right thing to say. We could instead say that she acts on those goals because they are the ones she accepts and is committed to, the ones that she thinks are genuinely valuable. Of course the goals that she accepts and is committed to now must be her present goals, but that does not necessarily mean that she acts on them because they are her present goals. Her attitude does not need to be construed as 'these goals are the appropriate ones for me to realize simply because they are my present goals, and a person's present goals as such always outweigh goals they may hold at other times'. So it is not clear that her stance does depend on giving greater importance to the present than to other times.[8]

Given the problems for the prudential solution, there is a more appealing way of understanding temporal neutrality. In present–future cases it does not require that the agent should give equal weight to his present and future values. He is only required not to give them unequal weight in virtue of some fact about timing. In other words, he should not give more weight to his present values because they are his present values. But temporal neutrality does not, for example, forbid him to give more weight to his present values on the grounds that they are in themselves more important or more reasonable than his future values.

[8] For this reason it is unfortunate that discussions of temporal neutrality often use examples of changing goals. We might agree that we should give equal importance to each temporal part of our lives (for example, we might think that a benefit in the present provides no stronger reason to act than a benefit of the same size in the future), and still be puzzled by cases of conflicting goals. Combining the two issues might make the requirement of temporal neutrality seem less persuasive than it actually is.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

We could revise prudence using this understanding of temporal neutrality. Then there would be no conflict with modest objectivism. My positive view will be compatible with this notion of temporal neutrality, but I will argue in Part V that this revised version of prudence needs to be supplemented by other ideas to properly handle cases of changing goals.

III

The second solution to the problem of conflicting values holds that a person should act on her present values in making decisions that will affect her life at later times when her values will be different. This answer is radically different from the first. It seems to give an importance to time itself that the first solution would reject on principle. It is undeniably focussed on the present in some sense, in that it claims that when we make a decision we should use the values that we hold at the time the decision is made.

This solution is defended by Parfit. He develops the view in opposition to the prudential solution, and he argues that it escapes the objections to the prudential solution. Parfit considers it as falling under a more general theory of reasons for acting which he calls the 'present-aim theory'. I will use this name, although I am interested in it specifically as a response to changes in values and not as a general account of reasons for action.[9]

There are two interesting features about Parfit's discussion. First, he treats the present-aim view as a theory about subjective rationality. In this respect he differs from Nagel who does not draw a distinction between subjective rationality and objective rationality. Second, Parfit narrows the range of cases that he is considering. He says that the points he wants to make are clearer if we consider desires or goals that rest on moral beliefs, and comments that Nagel's examples concern ideals about how we should live our own lives rather than explicit moral principles. A suitable example might be someone whose beliefs about justice change between youth and old age.[10] Later I will explain why Parfit's conclusions are persuasive in the moral cases but less persuasive in cases involving personal ideals.

[9] Parfit's aim is to make a case for the present-aim theory and show that there might be persuasive answers to the objections against it. However, I think that what he says about changing values specifically (Parfit, *Reasons*, pp. 153–6) suggests that he sees it as the best response to that problem, whatever its merits as a general account of reasons for action.

[10] Parfit discussed such an example in Derek Parfit, 'Later Selves and Moral Principles', *Philosophy and Personal Relations*, ed. A. Montefiore (London, 1973), sect. IV.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

Parfit thinks it is clear that the present-aim view does better than the prudential solution. He comments that the prudential solution asks an agent to give the same weight to values that she (currently) accepts and values that she (currently) rejects, and this is obviously an unreasonable demand to impose. Presumably his confidence in making this claim is explained by his decision to talk about what is subjectively rational.[11]

The present-aim view does not allow merely future goals (goals that we will only come to accept in the future) to have any bearing on present decisions in present–future examples. It also does not allow merely past goals (goals that we accepted in the past but later abandoned) to influence present decisions in past–present examples. Parfit emphasizes the latter point in contrasting the present-aim view with the prudential solution. This aspect of the view is indeed persuasive, but it also points to a feature of the theory that is less intuitively acceptable. In present–future examples the present-aim view does permit our present goals to influence our future lives at times when we will no longer endorse those goals. It tells the man in Nagel's example to act on his present desire to face dangers in the distant future, despite his knowledge that at that future time he will value safety and security. In fact the present-aim view is more radical than the prudential solution, since it says he should only consult his present goals while the prudential solution asks him to take into account both his present goals and his future goals. If he does act on the present-aim theory the result will be that his life in the future will be determined by goals that are merely past goals at that stage in his life. So the present-aim view permits merely past goals to have influenced the present, although the supposed unreasonableness of this was offered as a reason for preferring the present-aim theory to classical prudence.[12]

Like the prudential solution, the present-aim theory conflicts with modest objectivism. My goals at T1 about T2 might have some value,

[11] Parfit comments that it might be logically impossible to do what the prudential solution asks us to do (Parfit, *Reasons*, p. 155). Presumably the thought is that if we were to give equal weight to our present and future values this would be incompatible with the supposition that we now endorse our present values and reject the future values. I am not sure that this is a real problem for the prudential solution, so I will concentrate on the objection that it asks us to do something unreasonable.

[12] As Parfit explains, the present-aim view does not mandate which values and goals a person should have at a given time. If an agent held the appropriate set of goals, the present-aim view would reach the same conclusion as the prudential solution. Consequently in the examples I use against the present-aim view it would be possible to specify a set of goals such that if the agent held those goals the present-aim view would reach what I consider the intuitively plausible conclusion about the example. But this does not mean that the examples are harmless. The present-aim view says that in cases of changing goals the agent should act on her present goals whatever they are, so the view would generate implausible conclusions in other cases.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

but they might have much less value than my goals at T2 about T2. So it would not be reasonable to give the first set of goals dominant weight and the second set of goals no weight.

Given Parfit's account of the present-aim view, the obvious reply will be that the view is only intended to be a theory of subjective rationality. That was unconvincing as a defense of the prudential solution, but the present-aim view is much more persuasive as a theory of subjective rationality. If we understand 'x is subjectively rational for a' as 'a believes that x is rational', and we are willing to apply this view to a person at a particular time, then the present-aim view seems to be the obviously correct account of what it is subjectively rational for a person to choose at a time, including the agents in examples of changing goals.

However, being treated as an account of subjective rationality is not an unmixed advantage for the present-aim view. If the view is understood in this way its status as the correct theory of subjective rationality is a simple consequence of the definition of subjective rationality. This is not the source for the present-aim view that its proponents seem to envisage. It is no longer based on a view about the significance of time for rational choice that assigns a special importance to the present as opposed to the past and the future. The present is singled out because the issue of how we should reasonably respond to a change in values has been reduced to the question 'What would it be subjectively rational for the agent to choose at time T1?'. Arguably, we do not learn anything important about either rationality or time when we see that the answer to that question will depend on the goals that the agent believes to be valuable at T1.

Moreover, many of us will think that there is objective rationality as well as subjective rationality. We might suspect that the most important issue will be what it is objectively rational to do in cases of changing values. If we defend the present-aim view in the way that I have described then there is no reason to think it will have any role to play in the account of objective rationality.

Faced with these concerns someone might try to connect the present-aim view with objective rationality. It cannot be all there is to objective rationality – if my present goal is objectively worthless and my future goal has great objective value, it cannot be rational to act simply on the basis of my present goal. But perhaps the special importance of the present can be regarded as at least one factor in determining objective rationality. The fact that I am now committed to a certain goal might contribute to the objective rationality of my decision to pursue that goal, even if it can be outweighed by other factors that are also relevant to determining what is objectively rational. I am not sure whether this possibility is worth exploring, but the view that would result would bear

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

little resemblance to the present-aim view. It must be explained and defended in a different way. We cannot support this view by pointing out that an agent will at least believe that an action that furthers his current goals is rational.

IV

The third answer to the problem of changing goals resembles the second in one respect. It identifies a particular time such that it is the goals that one holds at that time which should determine decisions in cases of changes of goals. However, it does not single out the time at which the decision is made for this special status. Instead it picks out the time at which the relevant consequences of the decision will occur. It does not appeal to the goals the agent holds at T1 concerning T2 in the way that the second solution does (assuming that T1 is the present and the time at which the decision is made). Instead it appeals to the goals that the agent will hold later at T2 when the effects of the earlier decision will be realized.

About merely past goals the third solution agrees with the present-aim view that these goals should not influence a decision made at T1 about T1. However, it disagrees with the present-aim view about present–future cases. The second solution claims that merely future goals – goals concerning T2 that you will hold at T2 but do not hold at the earlier time T1 – should not influence a decision made at T1 about T2. The third solution takes the opposite view. It says that it is the goals that you will hold at T2 which should determine a decision made at T1 about your life at T2.

The present-aim view is naturally expressed in terms of tense, the distinction between past, present, and future. It holds that when a person makes a decision the values it is reasonable to use are the values that she accepts when the decision is made, and those values are her present values. Merely past values and merely future values are excluded. The third solution is most naturally expressed in a way that does not mention past, present, and future. In assessing a change in a person's life that occurs at T1 the values to be applied are the values that the person holds at T1. When we make decisions about T1 this principle will sometimes ask us to appeal to present values, sometimes to future values, and sometimes to past values (for example, when we are judging the significance of some event in our life in the past), depending on the relationship between T1 and the time at which our decision is made. I will call the third solution the 'simultaneous values view', because it maintains that the relevant values are the values that

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

the agent accepts simultaneously with the temporal part of his life that is being evaluated.[13]

Krister Bykvist defends a particular version of the third response.[14] He would argue that its conclusion about Nagel's example – the man should not compel himself to face dangers at some future time when he will be deeply averse to taking risks – is much more persuasive than that of the second solution. However, Bykvist's version of the view does not completely ignore our past preferences (hence the title of his *article*, 'The Moral Relevance of Past Preferences'). He thinks that a preference at T1 to be a poet at T2 has no weight if at T2 itself one will prefer to be a philosopher, but if the past preference is supported by one's current preference (that is, if the past preference and the present preference agree) then the past preference does matter for its own sake. So if at T2 I do still wish to be a poet my reasons for being a poet at T2 are stronger because I held this preference in the past as well as the present. Suppose that at T2 I come to have a brand new preference to be an athlete, which is equal in strength to my preference at T2 to be a poet. According to Bykvist I have stronger reasons at T2 to be an artist than to be an athlete, because the first present preference has the additional support of my past preference while the second present preference does not.

Like the other solutions, Bykvist's proposal conflicts with modest objectivism. The goals concerning T2 that I hold at T2 may be less valuable than the goals focussed on T2 that I held at T1, so a modest objectivist would not necessarily agree that my decisions about T2 should be based on the values I hold at T2. Bykvist sets objectivism aside. He assumes that my preferences at T1 and T2 cannot be criticized or rationally assessed, so that we cannot say that my goals at T1 are either more or less reasonable or valuable than my goals at T2.

[13] A fourth view should be mentioned for the sake of completeness. Instead of making a claim about the significance of the temporal distinction between past, present, and future, it makes a claim about the significance of the different temporal distinction between earlier and later (or prior and posterior). The claim is that in cases of conflicting goals the person should act in accordance with the goals that he will hold later in his life. This view would agree with the present-aim view and the simultaneous-values view about past–present cases, and it would agree with the simultaneous values view about the kind of present–future cases that I have discussed. However, we can imagine a different kind of present–future case. At T1 I must make a decision that will affect my life at T1, but I know that later in life at T2 I will hold different goals which I would then retroactively apply to my decision at T1. The simultaneous values view tells me to make the choice at T1 in accordance with my goals at T1, but this view would tell me to make this choice in accordance with the goals I will come to hold at T2. This conclusion is intuitively implausible, but it does not seem to me to be as implausible as the conclusion that I should act in accordance with my past goals in past–present cases.

[14] Krister Bykvist, 'The Moral Relevance of Past Preferences', *Time and Ethics: Essays at the Intersection*, ed. H. Dyke (Dordrecht, 2003).

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

Nevertheless, his solution to the problem of changing goals has the status of a view about objective rationality, not subjective rationality. He thinks that in this example it is rational for me to make my choice at T1 on the basis of the goals I will hold at T2, not the goals that I hold at T1. If instead I were to make the choice in accordance with my goals at T1 because I thought that was the reasonable thing to do Bykvist would say that I was making a mistake.[15]

Bykvist provides a rationale for the view. He says that a person is a self-identical entity that persists through time, but at any given time it is a temporal stage of that person that has preferences. A person can be seen as a related series of person-stages. Bykvist believes that a view of reasons for action should reflect both of these aspects of a person. A person-stage should have a say in what happens to it, but it is questionable that it should get a say in what happens to other person-stages. Allowing a person's past values to determine his present reasons for action would violate the autonomy of the person-stage that holds a set of values now, in the present. However, a view that claims that only a person's present reasons and goals matter for his present actions – as the present-aim theory does – is mistaken because it ignores the significance of the fact of the continuing self-identical person, and it ignores the fact that we live our lives partly from a diachronic perspective. So Bykvist presents his view as properly respecting both aspects of our lives in time. Because of the importance of a particular person-stage, present goals and values outweigh past or future goals and values when there is a conflict (again assuming the agent is assessing his life in the present). But past values should be credited with some importance in their own right when they match the person's present values.

In what he says about the importance and the autonomy of a person-stage Bykvist is arguably adopting a view about personal identity that some will challenge. And even if we do not object to his rationale in this way it is not clear that it justifies the conclusions that he draws. Suppose we grant that both the continuing person and the person-stage are significant, and that a view about reasons for action (and a solution to the problem of changing goals) should reflect the importance of both. We might think this gives us a reason to count present values as being more important than the values the person holds at other times, but Bykist gives them lexical priority. Values held at other times in my life only matter at all if they agree with my present values. If we think that

[15] Bykvist does not say whether my predictable future preferences about the present should have weight if they agree with my present preferences about the present, but the ideas in his view should lead to that conclusion.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

both the continuing person and the person-stage matter, why should we draw such a one-sided conclusion?[16]

I have criticized Bykvist's rationale. Nevertheless, the view itself seems to me more reasonable than either the prudential solution or the present-aim view. I say this because the simultaneous values view seems to better match the intuitive judgments that we would make about present–future and past–present cases. In particular in Nagel's example this view tells the man not to commit himself to dangerous adventures in the future. Both the prudential solution and the present-aim view have serious problems accounting for those judgments.

V

The positive view that I will defend resembles Bykvist's in that it reaches many of the same conclusions about the examples, conclusions that make Bykvist's view seem more persuasive than the alternatives I have already considered. However, it gives a very different explanation of those conclusions.

Unlike the views considered so far, this view is designed to be compatible with modest objectivism about goals and values. It also differs from them in another way. The three views I have already considered see the problem of changing goals as one about the relationship between time and rationality. To answer the problem they look for some attitude towards time that counts as rationally required. The prudential solution thinks that rationality requires temporal neutrality. We should see the different times in our lives as having equal importance, and on this basis we should conclude that the reasonable thing to do is to take both sets of goals into account. The present-aim view believes in the fundamental importance of the temporal divisions of past, present, and future. It holds that because of that importance a reasonable agent who makes a decision should use the goals that he endorses at that time, not goals that he held or will hold at other times but which now belong to the past or future. The simultaneous values view focusses not on the time at which the decision is made but on the time at which that decision will impact the agent's life. It proposes that the agent should decide using the goals he will have when his life is changed, not the goals he may endorse at the time of the decision that

[16] Bykvist's rationale spreads to cases that do not involve changes in goals. Suppose that a person must choose between a gain of well-being in the present and an equal-sized gain of well-being at some future time, in a case where there is no relevant change in the person's goals or values. The view about the significance of person stages seems to imply that the person should choose the present gain. Both choices would benefit the continuing person equally but only the first choice benefits the person stage that makes the choice.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

leads to the change. The last two views give a special importance to the present, though not in the same way. Or rather they identify a different present as being especially relevant to the decision.

The final view does not look for a response to the problem in some general attitude towards the significance of time. Instead it appeals to how we should understand welfare or well-being. In examples of the sort I have considered the agent will typically decide by thinking about which action will best promote her own welfare. She wants to make the decision that will make her life better on the whole (the broad temporal scope of that aim does not beg the question against the present-aim view).

My account of well-being will also involve objectivism. Just as I have assumed that some goals can be more important than others I will also assume that achieving a more important goal will better promote a person's well-being than achieving a less important goal.

However, I do not think well-being is simply determined by the objective value of states or activities. There is another relevant factor. Put with maximum generality, the agent should also positively respond to the valuable state or activity if well-being is to be maximized. There might be different views about the nature of this response. Some might see it as positively evaluating the activity, others as welcoming it or enjoying it, or as desiring the activity, or as some combination of these responses. And the response condition might be given different degrees of strength. Some might contend that anything that receives the positive response contributes to the agent's well-being at least to some extent regardless of its objective value, others might think that activities that do possess objective value will not contribute to well-being at all unless the agent positively responds to them in the appropriate way, and so on.[17]

For the purposes of this article it is not necessary to answer these questions. The important thing is that positive response is a factor in determining well-being just as objective value is. In what follows I will use a relatively weak account of its strength. I will assume that it influences well-being at least to the extent that a person might receive more well-being from a less valuable activity that they positively respond to than from a more valuable activity that they do not positively respond to.

The positive response condition applies to cases where values change. If at T1 I hold values that are strikingly different from the values that

[17] Recent work on well-being by Ronald Dworkin, *Sovereign Virtue* (Cambridge, Mass., 2000), ch. 6; Wayne Sumner, *Welfare, Happiness, and Ethics* (Oxford, 1996); and Stephen Darwall, *Welfare and Rational Care* (Princeton and Oxford, 2002) seems to accept some version of the positive response condition.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

I will hold at T2, and if I use my values at T1 to make a decision at T1 that will affect T2, then when T2 comes I will not positively respond to the change I have brought about. That would reduce the amount of well-being that I will experience at T2. Reaching this conclusion does not depend on supposing that my values at T2 are more reasonable than my values at T1. Whether I think about this question before or after the change in my values I will reach the same conclusion about the effect of the decision on whether the positive response condition will be satisfied at T2.[18]

An example will be helpful. Suppose Alice is a writer thinking about her old age when her active career will be over. She now believes that the best thing she can do in old age would be to preserve her legacy as a writer. She might advise younger writers, prepare revised editions of her books, lecture at university campuses, and so forth. But she also knows that when she is old her mind will have changed. She will then most want to engage in the kind of social activities that she neglected when she was younger and she will also come to think that this is the most worthwhile way to spend her final years. There is something she could do now that would lock herself in to the first kind of life in her old age. What choice should she make?

For the sake of argument let us suppose that her earlier view about value is the correct view. Social activities in retirement would have some value, but not as much as continued devotion to literature. Nevertheless, if the ideas I have been explaining are reasonable, this does not settle the question of what choice she should make, given that she is concerned with maximizing her own well-being. Literary activities would have more value than social ones, but if she chooses them the positive response condition will not be satisfied when she engages in them. If she makes a choice now that allows her to engage in social activities during retirement it will be satisfied. So it is possible that she will have a higher level of welfare in old age if her life is centred on the social activities, and she herself can appreciate this when she makes her decision.

The example shows how this view is compatible with modest objectivism. We can reach the conclusion that most people would find intuitively persuasive even though we admit that the literary activities would have more objective value than the social ones. Since the positive response condition is only one determinant of well-being, in another

[18] The same cannot be said about the objective value determinant of well-being. My judgment of the objective value of what happens at T2 will be different depending on whether I make this judgment at T1 or T2, so my view about the amount of well-being that I will experience at T2 will also differ depending on whether I am relying on my values at T1 or my values at T2.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

case the better choice might be to lock oneself in to activities that one would not appreciate at the time one in engaged in them.

The example also shows that the view about well-being can explain judgments that people might see as favouring one of the earlier theories I have discussed. In this case the view draws the same conclusion as Bykvist's view. But the conclusion does not depend on the blanket claim made by the simultaneous values view that the decision about her old age should be made simply on the strength of the values that she will hold when she is old. We have admitted that her views in old age about the respective values of literary and social activities are mistaken. If she thinks in terms of well-being she can decide against locking herself into a literary life even if she uses her present values in making the decision. There is a sense in which her values during old age do determine the decision that she should make earlier in her life, but not because they have a unique importance for her decision simply because they are the values that she will endorse when she is old. They matter because they determine whether the positive response condition on well-being will be met during her old age. Since my view accepts objectivism about goals while Bykvist's does not, they will reach different conclusions in other examples. By discussing changes in goals in terms of well-being I am not assuming the egoistic view that choosing rationally requires maximizing one's own long-term well-being. That theory of rationality is mistaken. However, one's own well-being is one goal among others that it is rational to aim at, although it might well be outweighed by others goals. In Nagel's example, and in my examples, the agent's concern is with living her life in the way that is best in terms of well-being.

I mentioned that Parfit moves away from examples about ideals applied to one's own life and considers examples about the moral treatment of other people. We can now understand the significance of this. In Nagel's examples the change in ideals can influence the positive response condition and so affect the agent's future well-being. But in the case of a moral commitment to a principle of justice, the fact that I will later endorse a false principle of justice does not mean that the unjust actions I will perform in following that mistaken principle will be less unjust because I accept that principle when I perform them. So Parfit was right to think that in the moral examples there is less reason for the agent to give weight to his future change of mind.

This view does not produce a solution to the problem by retreating to describing subjective rationality. The ideas in the view are part of a theory of objective rationality, although of course they do not amount to a complete theory of objective rationality. In my example it is objectively unreasonable for the writer to lock herself into a life of literary activity

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

in her old age. If she reached the conclusion that it was a rational thing to do she would be making a mistake.

The example of the writer suggests a qualification of the view. Perhaps some non-moral ideals should be treated in the way Parfit treats moral ideals. The writer might attach great value to literature and to activities associated with it, and let us again suppose that she is objectively right to do so. She might think that it is more important to maintain her commitment to a literary life for the sake of literature rather than for her own sake. She does not believe that this choice would best promote her own well-being, she believes she should live in this way because of the value of literature even though it diminishes her well-being. I would not argue that someone would be unreasonable to think in this way. It might make sense to sacrifice your own well-being for the sake of literature just as it makes sense to sacrifice your well-being for the sake of justice. But this is not the example that I have been discussing. In my example the agent is concerned with her own well-being, but she is not sure what to do given that her goals will change over time.

However, this variation on the original example can help us to choose between the view I have defended and the simultaneous values view. Suppose that the writer is concerned with aesthetic value, not with her own well-being. Still, she must face the problem that her current judgment of the aesthetic value of the literary activities possible in her old age differs radically from the judgment she will make when she is old. She now thinks that kind of postscript to her writing career would have considerable aesthetic value, but when she is old she will believe instead that such quasi-literary activities are merely pretentious. Bykvist's view would say that she should use the goals that she will hold in the future to make a choice that will affect her life at that future time. So it says that she should make her decision now in accordance with the aesthetic judgment she will make when she is old, even though we are now supposing that the reduction in her future well-being that would result if she used her present judgment instead is not a relevant consideration. It seems to me that this conclusion has very little intuitive plausibility, once the question of aesthetic value has been separated from effects on the agent's well-being. Why should we think it is irrational for her to choose on the basis of her current beliefs about aesthetic value? This suggests that when we find Bykvist's conclusions persuasive it is because we are being influenced by considerations about well-being.

A different example supports the same point. Suppose that I am making a choice at T1 that might bring about a change at T2, but I will never know at T2 whether this change actually does occur. My values at T1 say that the change would be a good thing while my values at T2 will

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

oppose the change. In this example my ignorance will prevent my well-being at T2 from being seriously affected by my choice at T1. Here too it is less plausible to say that my choice at T1 should be made according to the values I will hold at T2. My view, unlike the simultaneous values view, can explain why this should be so.[19]

VI

However, there is a problem for this response to the problem of changing goals. It might be true that at T2 I will have less well-being if my life fits my past values at T1 rather than the new values that I will hold at T2 itself. Suppose that for this reason I decide at T1 that my life at T2 should match my values at T2. Making this choice will presumably reduce my well-being at T1 (and at all subsequent times until the change in my values occurs) when I look forward to my life at T2. Why should we think that considering well-being will lead to the conclusion about the example that intuitively seems the right one, that I should be guided by the values that I will hold at T2?

This might seem to generate a fundamental objection to the view I have defended. In applying the positive response condition I have looked for a positive response on the part of the agent to some activity only at the particular time at which he performs the activity. This restriction on the temporal scope of the condition can seem unmotivated. If the agent positively responds to the activity at some other time – at some time before or after the activity is performed – why could this not contribute to his well-being? If I insist that only a positive response which is simultaneous with the activity counts as satisfying the condition it might seem that my view does after all depend on giving a special significance to a particular time, much like the simultaneous values view does. Then it would not be an alternative to views that pick out an attitude to time as being distinctively rational. It would make such a claim itself in how it interprets the positive response condition.

In considering this objection we should distinguish between two different ways in which the positive response condition can contribute to well-being. The first concerns the hedonic aspect of well-being. If at T1 I do not positively respond to my life at T1 this will cause me felt distress at T1 and lower my level of well-being at T1. It is clear that this effect can also occur across time. If in a present–future case I choose to act on my future values this will cause me distress in the present when I look forward to my future. And if I make a decision about my present life on the strength of my present values rather than my future values

[19] On some views of well-being if the change occurs it can affect my well-being at T2 despite my ignorance. Still, this example reduces the impact on well-being.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

I will feel distress in the future when I retrospectively look back at my choice.

I do not think this constitutes a strong objection to my view. We should take the pleasure and distress felt at other times into account, but I think it will typically be the case that the distress I feel in the present if my present life clashes with my present values will be greater than the prospective or retrospective distress in these examples. So I think the view will still be able to reach the conclusions about the examples that seem reasonable.

However, there is another way the agent's positive response might influence well-being. If I perform valuable activities the fact that I also positively respond to or approve of those activities might enhance the value of the activities and so contribute in that way to my well-being. In other words the activities will have more value (and a kind of value that contributes to well-being) in virtue of my positive response than if I had performed them without positively responding to them and without believing that they were valuable and worthwhile. I will call this the 'value enhancement' reading of the positive response condition.

Obviously this is a more challenging, and more controversial, claim about positive response. Some writers about well-being will reject it, but others will accept it.[20] I will not try to defend the claim but I will consider whether it is reasonable to broaden its temporal scope.

The objection to my view claimed that it was arbitrary to restrict the positive response condition to a positive response in the present to an activity performed in the present. I may not positively respond to a valuable activity when I engage in it, but if at a later time I retrospectively endorse and appreciate its value some might hold that this can also influence the value of the activity and the quality of my life.

One difference between the value enhancement version and the hedonic version of the positive response condition is that in the case of hedonic goods the increase of my well-being would be temporally located at the later time from which I look back at the activity and enjoy it. But if we believe in retrospective value enhancement it seems more natural to locate the increase *temporally* in well-being at the earlier time at which the valuable activity was performed.[21] Retrospective approval at T2 enhances the quality of my life at T1 itself, in the sense that it makes my life at T1 better than it would have been if I had

[20] Of the writers I have mentioned I think that Dworkin and Darwall would accept the enhancement claim, although they might express it in different language. Sumner would not accept it.

[21] Another possibility is that the gain in well-being cannot be assigned to any particular time in the agent's life. However, it seems appropriate to associate the gain with the activities that the agent is responding to.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

never at any time in my life appreciated the value of the activity that I performed at T1.

Retrospective value enhancement seems easier to accept than prospective value enhancement. In present–future cases I do not think we would say that my present prospective approval of a future life that matches my present values but will clash with my future values provides a reason based on well-being to lock myself into such a future. We will simply think that when the future arrives the positive response condition will not be satisfied, and we will draw the appropriate conclusion about my future well-being. So prospective value enhancement is not a threat to my view.

However, I think that retrospective approval has more significance. If at the end of my life I am fortunate enough to come to see the value in things I accomplished in the past when I did not take them seriously this might in some sense redeem the value of my past. Suppose an artist created excellent paintings during his career, but did not at the time in any way appreciate or value what he was doing. He produced the work just for the sake of money and prestige. But in his later years he begins to value his paintings and his own talent for what they were worth. For example, he now understands and relishes the aesthetic appropriateness of the choices that he made in his paintings, choices that came out of him carelessly and instinctively when he was young. This belated recognition does not make the pictures themselves better or worse. But I think it might influence our judgment about his well-being and how good a life he lived. It means that he is at least not permanently alienated from his own most important accomplishments.[22]

Nevertheless, even in this example I do not think that retrospective approval gives his activities as much value as they would have had if he had positively responded to them while he was engaged in them. And as long as appreciation of the activity when it is performed has a greater effect on value than retrospective appreciation then I think the point I wanted to make still stands. Well-being provides a reason for using the values I will hold at T1 when I make a decision that will

[22] In a different example a person might look back at his life and realize – correctly – that his past activities had less value than he believed when he engaged in them. Does this kind of retrospective disapproval affect his well-being, and is the effect positive or negative? Here I am tempted to appeal to the thought that the positive response should be proportional to the value of what it is a response to. His life is better because he does at last respond in the appropriate way, though it is not as good as it would have been if his past beliefs about the value of those activities *had been true rather than false*.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

affect my life at T1, and so it supports an answer to the problem of changing goals.[23]

Proposing an asymmetry between prospective and retrospective approval might itself seem to undermine my account of the foundations of this view. I have said that the basic claim is about well-being, not about the significance of time. However, perhaps prospective approval does not enhance value because the relevant values belong to the past when the valuable activity is performed. Retrospective approval is more effectual because, although the person does not already endorse his future values when the activity is performed, at least those values have not already been rejected and consigned to the past. If so the explanatory work might be done by a claim about the respective importance of the past and the future.

This diagnosis might be correct. In any case it is undeniable that the view gives importance to facts about time in determining what it is rational to do in cases of changing goals. Nevertheless, if the asymmetry holds it will be true that retrospective approval contributes to well-being in a way that prospective approval does not, and this is why it is reasonable to make a particular choice. I think this makes it fair to characterize the view as being concerned with well-being. The significance of facts about time for the rationality of the choice is mediated by a particular account of well-being. The view does not make a direct claim about the rational importance of some relationship between the time at which the choice is made, the time of events being evaluated, and the goals the agent holds at those two times.

I have suggested that in cases of changing values the rational choice will (at least often) be the choice that would maximize the agent's well-being, assuming that the theory of well-being involves the positive response condition. It may seem that this is to offer a superficial answer to a deep question. Some will feel that a proper answer has to begin from some fundamental stance about the rational significance of time, like the stances that motivate the prudential solution, the present-aim view, and the simultaneous values view.

A point I have repeatedly made is that it is hard to make these stances fit with modest objectivism. We can recast the views as views about subjective rationality. If we think that there is such a thing as objective rationality, then it is arguable that the views will not be answering the most important question about changes in goals – what

[23] Accepting retrospective value enhancement means that in some cases it would be better for me in terms of well-being to perform *now* more valuable activities that conform to my future values rather than less valuable activities that conform to my present values. But I would contend that this is not an implausible conclusion to reach about those examples.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press

would it be objectively rational to choose? If we think it is pointless to look for more than an account of what it would be subjectively rational to choose we are likely to arrive at the present-aim view. However, we are committed to a very strong thesis about rationality, and our conclusions about changes in values are completely explained by this thesis.

A modest conclusion for this article would be that the points made about well-being are relevant to decisions made in the face of conflicts of values over time, and they have not been taken into account. The immodest conclusion would be that most discussions have approached the issue in the wrong way. Each of the positions about the significance of time faces persuasive counter-examples because all those positions are false. In cases where a person's values conflict over time reason does not tell us to rely on our present values, or to rely on our future values, or to strike a balance between present values and future values. These stances have seemed reasonable to some writers because they have accepted questionable views about the nature of reasons, or because they have misused the distinction between objective rationality and subjective rationality, or because they have drawn conclusions favourable to the stances from examples that are better explained by a different consideration.[24]

mckerlie@ucalgary.ca

[24] I would like to thank Nils Holtug, Krister Bykvist, Thomas Hurka, and audiences at the University of Calgary and the University of Toronto for helpful comments on earlier versions of this article.

https://doi.org/10.1017/S0953820806002342 Published online by Cambridge University Press