Comments on Dennis McKerlie's 'Rational Choice, Changes in Values over Time, and Well-Being'

KRISTER BYKVIST
Jesus College, Oxford

I am grateful to have been given the opportunity to comment on McKerlie's interesting article, especially since it concerns one of my pet topics and provides many helpful comments on one of my own articles on this topic. My comments will be brief because I agree with most of his points, in particular, his criticisms of the prudential view and the present-aim theory. What I want to do here is just to clarify a couple of things concerning my own theory, concede some of the difficulties that McKerlie raises for my theory, and see to what extent his own proposal fares better than my own theory.

As McKerlie points out, the theory he wants to defend shares many of the implications of my own theory. But he claims that his theory gives a different and more plausible explanation of these implications. In my article, I argued for the idea that the problem of past preferences has nothing to do with pastness as such. This I still believe is true and McKerlie is on my side here. What I am not so sure about is whether the theoretical underpinnings I gave to my theory are convincing. In my article, I argued that the problem of past preferences arises when these preferences of past selves 'poke their nose' into present selves' lives and private concerns. This way of putting things suggests that what is at issue is some kind of respect for the autonomy of person-stages, and I chose to spell out this respect in the following way: each person-stage has a veto over what they should do with their life-stage so any conflicting preferences of other persons or other temporal stages of the same person are to be completely disregarded.[1] I stressed, however, that in the intrapersonal case the autonomy of person-stages is not the only thing that matters. Since we do not just lead our lives from the perspectives of individual moments but also from a more comprehensive diachronic point of view, it seems plausible to count those past preferences of a person that agree with her present

[1] K. Bykvist, 'The Moral Relevance of Past Preferences', *Time and Ethics: Essays at the Intersection*, ed. D. Dyke (Dordrecht, 2003), p. 116, and p. 127.

© 2007 Cambridge University Press
doi:10.1017/S0953820806002354
Utilitas Vol. 19, No. 1, March 2007
Printed in the United Kingdom

https://doi.org/10.1017/S0953820806002354 Published online by Cambridge University Press

preferences. From the diachronic perspective, it is important that our lives show a unity and continuity of purpose so that preferences at one time link up with preferences at a later time.[2]

McKerlie rightly points out that my talk about person-stages suggests a controversial theory of personal identity, but my intention was not to take a stand on this thorny metaphysical issue.[3] My use of 'person-stage' was only meant to be a convenient way of referring to a person as he is at a particular time (or during a certain period).

McKerlie also claims that my theory is a version of the 'simultaneous values view', according to which 'the relevant values are the values that the agent accepts simultaneously with the temporal part of his life that is being evaluated'.[4] This is misleading since it suggests that I only want to give weight to synchronic preferences. But, as McKerlie himself acknowledges, I do give weight to diachronic preferences concerning a certain time if they are supported by the synchronic preferences at that time. McKerlie is also mistaken when he claims that my theory gives a special importance to the present.[5] My theory is time-neutral in the sense that it is irrelevant whether a preference is past, present, or future. What is relevant according to my theory is whether a diachronic preference for a certain time is supported by the synchronic preferences at that time.

More important is McKerlie's objection that I give lexical weight to synchronic preferences, the preferences a person-stage has towards its own period. He points out that nothing in my theoretical underpinnings warrants this one-sided conclusion, for it is one thing to give priority to synchronic preferences because we should respect the autonomy of person-stages, quite another to give them lexical weight.[6]

The reason why I chose to give lexical weight to synchronic preferences was that I wanted to avoid the present-for-past sacrifices illustrated by Parfit's poet-example. If past diachronic preferences are given some significant weight even in cases where they conflict with the corresponding present synchronic preferences, we would have to accept that a person may have reason to sacrifice her present preferences about her present life ('not to be a poet') for the sake of her past preferences about her present life ('to be a poet').

McKerlie's own theory accepts present-for-past sacrifices when the objects of the past preferences are much more objectively valuable than the objects of the present ones. I agree with this and tried to make this

[2] Bykvist, 'Moral Relevance', p. 127.
[3] McKerlie, p. 62.
[4] McKerlie, pp. 60–1.
[5] McKerlie, p. 62.
[6] McKerlie, p. 62.

https://doi.org/10.1017/S0953820806002354 Published online by Cambridge University Press

clear in my article. So, McKerlie is mistaken when he claims that my account conflicts with this kind of modest objectivism.[7] What I claimed in my article was that my theory should come into play when we have weeded out those preferences that have been formed in a suspect way and those that are intrinsically irrational, i.e. concerned with objects that lack objective value.[8] Obviously then, my theory will not conflict with objectivism. My theory only tries to find a solution to the problem when the conflicting past and present preferences concern objects that do not differ in objective value.

So what shall we do when our past and present preferences conflict but concern objects that do not differ in objective value? Whereas my theory avoids present-for-past sacrifices by giving priority to the autonomy of person-stages, McKerlie's avoids these sacrifices by invoking a positive response condition on *well-being*: objectively valuable activities and achievements will have more value for a person if she positively responds to them in some way.[9] He adds the important qualification that only simultaneous or retrospective endorsements can enhance well-being.[10] Prospective endorsements of future lives will not enhance well-being, at least if these lives will not be endorsed in the future. This asymmetric treatment of the endorsement condition guarantees that if my present self does not endorse the activities my past self insisted he should take part in, there is no reason based on well-being to satisfy my past prospective preferences.

I agree with McKerlie that invoking an endorsement constraint on well-being might be a promising way to deal with the problem of past preferences. But I doubt that the qualifications he adds are plausible. First of all, if McKerlie wanted to ban *all* prospective preferences, his theory would be too restrictive. Surely, adopting some worthwhile ambitions and pursuing them with great interest until finally they are fulfilled will make a significant contribution to a person's well-being, but this will involve the satisfaction of prospective preferences.

Second, his theory would also be too restrictive if it only banned prospective preferences that will later lack support. Suppose that I have been working all my life to publish a book. Right before I send it off to the publisher I die but, fortunately, one of my friends, knowing how much this book project meant to me, sends the manuscript off to the publisher and the book is later published. (If you find the notion of posthumous benefits problematic you can imagine that instead of dying I fall into a coma.) I don't see why this action could not be said to make

[7] McKerlie, p. 61 and p. 63.
[8] Bykvist, 'Moral Relevance', p. 130.
[9] McKerlie, p. 64.
[10] McKerlie, pp. 69–71.

https://doi.org/10.1017/S0953820806002354 Published online by Cambridge University Press

my life better for me in virtue of realizing one of my most important goals. But a ban on unsupported prospective preferences would rule out this judgment.

Third, McKerlie assumes that retrospective endorsements enhance the well-being of *my earlier self*. So, my retrospective approval at T2 of my life at T1 enhances the quality of my life at T1.[11] But this will have very counterintuitive consequences. To use a version of McKerlie's own examples, suppose that you created excellent paintings in your youth but, on the whole, loathed the creative process.[12] You did it only for the money. In later years when you look back you may appreciate what you did and endorse that earlier part of your life. McKerlie has to say that you had a good life in your youth despite the fact that you hated what you were doing at the time. Of course, we do not need to deny that your earlier life was good in the sense of being a good *artistic* life. But to say that your former self is made better off retroactively by your later self's endorsement would take away what we find so tragic about your earlier life. You had it all in terms of objective value but were unable to endorse it. It is more sensible to say that the daily well-being you enjoyed in your youth is a thing of the past, and there is nothing we can now do to change this.

It is still true, as McKerlie points out, that retrospective endorsement can redeem the past. But the endorsement redeems the past not in the sense of making your former self better off but by making your life *as a whole* better for you.[13] It is a mistake to think that what is good for a person must always show up at some particular time. I would therefore want to suggest that my retrospective approval at T2 of my life at T1 can only enhance the value of my life as a whole. I see no reason why we should not also say that *prospective* endorsements can enhance the value of a life as a whole. The only argument McKerlie raises against this possibility is that satisfying a prospective preference will not affect *future* well-being if in the future the positive response condition is not satisfied. But this would create an odd asymmetry: retrospective endorsements can enhance past well-being, but prospective endorsements cannot enhance future well-being. Indeed, since McKerlie himself argues that the answer to the problem of past preference should not hinge on temporal issues, a better version of his theory would deny this asymmetry and instead allow that both prospective and retrospective endorsements can enhance the value of a life as a whole, but not the value of a life at a particular time.

[11] McKerlie, p. 69.
[12] McKerlie, p. 70.
[13] Here I follow D. Velleman, 'Well-Being and Time', *Pacific Philosophical Journal* 72.1 (1991).

https://doi.org/10.1017/S0953820806002354 Published online by Cambridge University Press

Exactly what implications this will have for the problem of present-for-past sacrifices depends on how much weight we want give to the well-being of an agent. As McKerlie concedes, it is implausible to think that the lifetime well-being of an agent is the only thing that matters for what she has most reason to do, since that would amount to some form of egoism.[14] Other factors can also come into play. This means that the autonomy of person-stages may turn out to have some significance after all. Rather than providing competing theories, perhaps McKerlie and I are best seen as providing partial theories, each of which identifies a valid reason for action.

[14] McKerlie, p. 66.

https://doi.org/10.1017/S0953820806002354 Published online by Cambridge University Press