Introduction

SERENA OLSARETTI

In a number of debates in contemporary moral and political philosophy and philosophy of economics, philosophers hold the conviction that preferences have normative significance. A central assumption that underlies this conviction is that a cogent account of preference-formation can be developed. This is particularly evident in debates about well-being. Those who defend subjective accounts of well-being, on which a person's life goes better for her to the extent that her preferences are satisfied, often qualify that account so that it does not include malformed or adaptive preferences (that is, preferences formed in non-autonomous ways, or humble preferences tailored to stifling circumstances), the satisfaction of which does not seem to contribute to well-being. This assumes that there is a normative standard of preference-formation with which to identify those preference that are malformed or adaptive in the relevant sense. An account of preference-formation is also important for philosophers who uphold an objective theory of well-being, on which well-being consists of the pursuit of objectively valuable goals, but who also believe in the value of freedom and thus emphasise the importance of respecting individuals' choices among various goals. For they, too, in extolling the importance of respect for choices, assume that these choices are not distorted by inauthentically formed preferences, and thus also need an account of preference-formation to help distinguish those cases in which we ought to respect people's choices from those in which we do not. In the vast literature on preferences, however, relatively little attention has been devoted to the explicit discussion of candidate accounts of preference-formation. The aim of the 2004 annual Royal Institute of Philosophy conference, which constitutes the basis for the present volume, was to foster such a discussion.

The papers collected here can, broadly speaking, be seen to address three main sets of questions on the topic of preference-formation and well-being. The first main set of questions concern the formulation of normative and descriptive accounts of preference-formation. The former are especially important, as they provide a standard by which we can discriminate between

https://doi.org/10.1017/S1358246106059017 Published online by Cambridge University Press

preferences that do, and those that do not have, prudential value. The papers by Richard Arneson, Connie Rosati, Johan Brännmark, and Mozaffar Qizilbash all start with the two-fold assumption that the satisfaction of preferences has some normative significance, where this significance is thought of specifically in terms of the relevance of preference satisfaction for our well-being; and that this significance varies depending on what the preferences in question are, or how they have been formed, so that it is possible that with respect to some preferences their satisfaction has little or indeed no prudential value. These papers then seek a standard with which we can discriminate between the prudential values of different preferences.

Arneson's "Desire Formation and Human Good" considers the three main families of theories of well-being or the human good—the desire or preference satisfaction account, the objective list account, and the hybrid account—with a view to identifying and examining their implications about which preferences it would be good for individuals to form. Arneson argues that the desire satisfaction account is committed to the problematic implication that people should form the desires that they can satisfy most easily, and that any attempt to avoid this implication from within the premises of the desire satisfaction account is unsuccessful. Ultimately, Arneson suggests, what underpins our unease with this implication of the desire satisfaction account is our implicit evaluation of some desires as more objectively valuable than others, an evaluation that requires the abandonment of the desire satisfaction account. The objective list account, and the hybrid account, fare better than the desire satisfaction account, in that they have the resources to offer a discriminating view of what desires people should form: both accounts suggest that individuals should form desires for goods that are objectively valuable. However, Arneson suggests that there is a tension between two different claims these accounts make. They may imply either that it is good for individuals to form desires that track what is of value, and thus reflect proper appreciation of it; or that it is good for individuals to form desires that make the pursuit of the good *effective*. These two things are not the same. For example, it may be that a disproportionately positive evaluation of the value of a particular goal one sets oneself—say, that of succeeding in one's career—can more effectively lead a person to achieve that valuable goal than a fitting and more moderate appreciation of its value. Arneson suggests that objective list theorists should adopt the second, rather than the first, view of what preferences people should form. Hybrid

https://doi.org/10.1017/S1358246106059017 Published online by Cambridge University Press

accounts of the good life, by contrast, will also emphasise the importance of there being a proper fit between people's preferences and the objects of those preferences, since on these accounts the *attitudes* that people have towards objectively valuable goods is an essential component of the good life.

Like Arneson, Qizilbash is interested in teasing out the implications of the central competing accounts of well-being for the formulation of a standard for discriminating between preferences. Moreover, he, too, doubts the capacity of purely subjective accounts to provide such a standard. Qizilbash proposes that a standard for discriminating between preferences should be able to identify adaptive preferences as problematic from the point of view of well-being, without positing an overly demanding set of conditions that preferences must meet in order to have prudential value. In particular, these conditions should reflect human limitations in forming preferences, including, for example, the constrained capacity to gather and appreciate information about the various options we may face. This limitation is ignored by desire satisfaction accounts which hold that the only desires whose satisfaction contribute to a person's well-being are those that person would have if she were fully informed about all the alternative options she faces.

In his "Leading a Life of One's Own: On Well-Being and Narrative Autonomy", Brännmark is also concerned that the standard for identifying which preferences have prudential value not be an overly demanding one. His suggestion is that we adopt the criterion of narrative autonomy as such a standard. Narrative autonomy is the autonomy an individual enjoys when she lives a life of her own, and is both protagonist and author of her own life. She must be a protagonist of her own life, as opposed to merely playing the role of a supporting character, and she must be author of her life, in the sense of having a certain independence of judgement about what she wants. Those preferences that hinder our leading lives that exhibit narrative autonomy, then, are problematic from the point of view of well-being. Narrative autonomy, according to Brännmark, constitutes a relatively undemanding standard of prudential value which can be satisfied by people with very different plans of life, and is more plausible than competing standards for discriminating between preferences, such as a standard that deems only idealised preferences which survive critical reflection as prudentially valuable. An important implication of Brännmark's argument, like Arneson's, is that we should abandon subjectivism about well-being, since arguably the theory

https://doi.org/10.1017/S1358246106059017 Published online by Cambridge University Press

of well-being within which narrative autonomy can play a role is an objective one that identifies the narrative meaning of lives as objectively valuable.

In contrast with the objective list and hybrid accounts examined by Arneson and Qizilbash, Rosati suggests that we discriminate between preferences by looking, not to the value of the objects of preferences, but to whether those preferences would be formed by individuals who have the capacity to form preferences well. To be good formers of preferences, individuals must have certain dispositions, which are best thought of as those they develop when successfully parented. Successful parenting, on Rosati's view, will exhibit, first, a concern for the agent-neutral value of children that fosters a robust sense of self-worth; second, a regard for the child as an autonomous agent, which fosters the capacity to respond to reasons, deliberate on desires, and engage in self-reflection; and third, a regard to the individuality of the child. In short, then, Rosati suggests that good formers of preferences have the dispositions that successful parenting promotes in children, and good preferences are those which we would form if we were ourselves guided by the regard towards ourselves which successful parents are guided by. Rosati's proposal casts new light on why adaptive preferences are sometimes problematic. For Rosati, adaptive preferences are problematic because the individuals who form them lack the dispositions that are necessary for them to be good formers of preferences. She can thus account for what seems deeply troubling about the paradigmatic cases of adaptive preferences, namely the adaptive preferences of "the hopeless beggar" or "the dominated housewife".

Finally, turning from normative to descriptive accounts of preference-formation, Daniel Hausman's paper argues that economists offer an account of preference-formation despite their claiming not to do so. More specifically, Hausman's main contention is that game theory can be seen to offer an account of how some preferences (namely, preferences over strategies) are formed, taking people's preferences over outcomes as given, and that moreover, in those cases where people's preferences over outcomes are affected by features of the game form other than its results (such as, for example, how a certain result is brought about), then even preferences over outcomes are not just given, and economists tacitly construct them. The claim economists often make, namely, that they do not and should not have anything to say about preference-formation, is therefore unjustified.

https://doi.org/10.1017/S1358246106059017 Published online by Cambridge University Press

The second main set of issues explored by papers in this volume concerns whether preferences conform with requirements of rationality, and what types of reasons we have to prefer certain things rather than others. In "Preference, Deliberation and Satisfaction", Philip Pettit defends a "value-serving" view of deliberation, on which the aim of deliberation is to find criteria with which to assess our goal-pursuing activities. On this view of deliberation, we deliberate not just *from* goals or desires, but also *about* them; we form our desires or goals in light of the desirability features of their objects, where these features could include the extent to which our desires cohere with one another, or the degree to which we are likely to be able to satisfy them. Desires are not just cravings which must be taken as givens, but are the objects of deliberation, and the significance of satisfying them, both from the first-person and the third-person perspective, will be a function of the desirability features of what is desired.

Christian Piller's paper also allows for the fact that the desirability features of the objects of desire may determine whether we have reasons to desire those objects. But his main concern is with whether we may be said to have reasons to desire something that track, not the desirability features *of the objects of desire*, but rather, *of the having of the desire* itself. In Piller's language, the question is whether we have "attitude-related", as well as "content-related", reasons to prefer or desire something, that is, whether the fact that having a desire for something makes it more likely that we will achieve something that is good for us is a reason to desire that thing (so that, for example, if preferring a saucer of mud over a pot of gold ensures that you will receive two pots of gold, you would have a reason to prefer a saucer of mud). Piller's answer to this question is affirmative for desires (while granting that the answer may be negative for other attitudes, like beliefs). If Piller is right, it can be rational to have preferences for what is (in terms of the desirability features of the objects of those preferences) worse, because to have such preferences can be better for us.

Even if we assume that there are requirements of rationality that our preferences must satisfy, there is a further question as to how we come to meet these requirements. In his contribution to this volume, John Broome asks whether we can come to meet those requirements by reasoning our way to them. Such reasoning could either start from the requirements themselves as premises (as the "second-order model" supposes), or proceed by considering the content of one's preferences (as suggested by the "first-order

https://doi.org/10.1017/S1358246106059017 Published online by Cambridge University Press

model"). Broome argues that the first-order rather than the second-order model of reasoning is more apt as an account of reasoning with preferences, where these are narrowly understood as desires, rather than, broadly, as dispositions to choose.

The third and final question addressed by papers in this volume concerns the normative significance of preferences that do not meet requirements of rationality or which do not seem to have prudential value, and, in particular, the significance of such preferences for the purposes of designing various policies (such as policies concerning what goods and services should be made available to people). Here the arguments by Robert Sugden on the one hand, and Cass Sunstein and Richard Thaler on the other, point in opposite directions. Both papers cite recent research in psychology and behavioural economics that points to systematic evidence of people's preferences being arational or irrational. But while Sugden believes that, in spite of this, a case can still be made for the presumption in favour of respecting people's choices, even when these reflect unconsidered preferences (preferences that do not meet the relevant requirements of rationality), Sunstein and Thaler suggest that there is a sense in which talk of the presumption in favour of "respecting" these preferences is altogether misleading here. More precisely, Sunstein and Thaler show that people's preferences are shaped by the context of choice people face, so that the setting up of that context determines what preferences people form. So, for example, there is evidence that more employees choose to commit to a savings plan if they are automatically enrolled into one and are given the option to opt out than if they have to opt in (where the savings plans are otherwise identical), thereby displaying a status-quo bias. This means that any policy inevitably affects people's preferences. Once we recognise this, it seems reasonable to suggest that we should so set up the context of choice as to promote preferences that it is rational or good for people to have, rather than setting up the context of choice in a way that leads people to have irrational preferences, or preferences that do not promote their well-being. Sunstein and Thaler grant that so setting up the context of choice is paternalistic, but insist that paternalism is inevitable, and that the paternalism they favour is *liberal* paternalism, in that it does not act by removing options, but by making an unavoidable decision concerning *which* options people should be offered.

The last paper in this volume, by Alex Voorhoeve, points to a problem with taking preferences as the relevant standard for policy that arises with both unconsidered and considered preferences.

https://doi.org/10.1017/S1358246106059017 Published online by Cambridge University Press

Voorhoeve's concern is specifically with welfarism, understood as the view that we should take the degree to which people's preferences are satisfied (as opposed to just respected, or not interfered with, which Sugden and Sustein and Thaler focus on) as the standard by which to evaluate people's situations for the purposes of policy. Voorhoeve points out that people's preferences, including their preferences about what preferences they have, change over time. For example, someone may now prefer to be, in the future, a struggling artist with artist's preferences over a successful banker with banker's preferences, but, if he becomes a banker and his preferences change, he will prefer to be a banker with banker's preferences. There is then an irresolvable tension between satisfying people's unchanged preferences on the hand and their changed ones on the other, in light of which, Voorhoeve suggests, it is desirable to abandon preference-satisfaction as a standard for policy.

Although a number of the papers in this volume are sceptical about allowing preferences to play too large a role in determining how it is prudent or rational for us to act, and what we owe to others, they all agree that respecting and satisfying preferences will play *some* role in how well a life goes, and in deliberation both from the first-person and the third-person perspective. If we share this starting point, the discussion of how preferences should be formed in order for them to have normative significance, which the papers in this volume make a contribution to, is of fundamental importance.

Acknowledgements

Acknowledgements are owed to the Royal Institute of Philosophy for sponsoring the conference that was the basis for this volume. Thanks also to the Analysis Committee and the Faculty of Philosophy, Cambridge, for further financial support, to St. John's College, Cambridge, for hosting the conference, and to Ross Harrison and Katherine Harloe for their contribution in organising the conference.