BIOGRAPHICAL INFORMATION

Benjamin Lange is a D.Phil. Philosophy candidate at the University of Oxford. His research focuses on issues in ethics and political philosophy. Currently, he is most interested in the ethics of partiality, in particular the ethics of parenthood, meaning in life, permissions and special duties, and agent-relativity.

doi:10.1017/S026626711800055X

A Philosophy for the Science of Well-Being, Anna Alexandrova. Oxford University Press, 2017, xlv + 196 pages.

Philosophers interested in the science of well-being come from an array of sub-specialties. This diversity of 'philosophies' speaks to the diversity of contexts in which well-being is an object of science. Diversity can be enriching, but it can also be philosophically distracting, particularly when we lack a framework to situate different perspectives. Until recently, the science of well-being has lacked such a framework. Anna Alexandrova's *A Philosophy for the Science of Well-Being* (Oxford University Press) provides this vital function – and more.

In this monograph, Alexandrova brings together a host of terms, e.g. 'value-aptness', 'well-being variantism' and 'evidentiary subjectivism', which together provide a conceptual space to locate the work philosophers do on the science of well-being. Alexandrova also puts forward a positive argument, that urges us to see well-being science as driven by context and normative values, while maintaining its objectivity and measurability. Even if one objects to the details of her argument, the terms she deploys to make her case will have a lasting impact on philosophers in this field.

Alexandrova's project is ambitious. The central question is, 'How can the science of well-being produce knowledge that is properly about well-being?' This is the Question of Value-Aptness. It encompasses three different issues:

(1) How should well-being be defined in a given scientific project?
(2) How should well-being be measured?
(3) How can the science of well-being retain objectivity?

The ambition of the project should now be partially clear. Alexandrova takes on big questions that have concerned rather distinct areas in philosophy: philosophy of well-being, philosophy of measurement, and the science and values literature. The science of well-being is not for the philosophically faint of heart.

In Part I (Chapters 1–3), Alexandrova defends a view she coins 'Well-Being Variantism' (WBV) according to which there is no single concept or theory of well-being that can do all the work we need it to do in all the different contexts in which it is used. This view largely aims to answer the first part of the Question of Value-Aptness. Part II (Chapters 4–6) addresses the second and third parts of Value-Aptness, and thus largely deals with matters in the philosophy of science.

The first two chapters of this book are the heart of Alexandrova's defence of WBV. She begins by noticing some practical realities. When philosophers discuss well-being, they tend to focus on general, all things considered well-being. When scientists use well-being they often focus on context-specific well-being, e.g. upper limb functioning. How should we understand this diversity?

Relying on contemporary epistemology and philosophy of language, Alexandrova puts forward three interpretations: Circumscription, Differential Realization and Contexualism. Arguing against the first two, she champions a version of the latter. Contextualism about well-being takes the diversity we witness to be a function of at least a partial change in the meaning of the term due to the context in which it is used. The context is at least partially determined by the features of the practical environment when an evaluation of well-being is made, e.g. the relationship of the parties involved, relevant contrast classes, etc.

For Alexandrova, well-being can mean different things in different contexts. But what is the content of these meanings? How should well-being be justified in science? Alexandrova's answer is that for methodological purposes, i.e. for justifying the concept as it is used in well-being science, we should adopt a position of theory diversity. Thus, we should not expect a single master theory to justify well-being in science. To make this argument she leans on Nancy Cartwright's work on scientific theories (Cartwright *et al*. 1995; Cartwright 1999).

Cartwright and colleagues use the analogies of scientific theories as vending machines and toolboxes. On the vending machine view, science is mechanized and theory is maximally important. When one inputs the initial conditions of any situation into the vending machine out pops a model of that situation. Cartwright argues that this view ignores the realities of scientific activity, especially the approximations and creative activity involved in scientific modelling. Alternatively, the toolbox view includes a multitude of tools that scientists can use to create models.

Adapting these analogies for the science of well-being, Alexandrova suggests that the vending machine view is similar to the view that a single, all-purpose theory of well-being can justify the different uses of this concept in this science. She uses an example, which although helpful is perhaps not as convincing as it might be. Her example is the QUALEFFO, a quality of life measure for patients with vertebral

fractures. If the vending machine view is going to work for well-being science, then according to Alexandrova, we have to be able to imagine combining a master theory, say idealized subjectivism, with assumptions about the population of people targeted by the measure, and end up with the well-being of this population 'conceived approximately as the QUALEFFO indicates' (37). Alexandrova hopes this proposal strikes the reader as far-fetched because as she explains, we do not have the bridge principles to move from the ideal theory to the items in the questionnaire. But I am not convinced it will strike every reader as outlandish a proposal as it truly is. To understand how bizarre the vending machine view is when applied to QUALEFFO, readers would need some insight into the decision points and complications of constructing such a measure. Cartwright and colleagues provide this sort of insight with detailed examples of, for example, superconductivity. Similar detail is missing in this book.

In place of the vending machine, Alexandrova suggests a toolbox. The aim of the toolbox is to justify the concept of well-being used in a particular research context. The tools will include different abstract or 'high' level theories of well-being such as hedonism and eudaimonism, and presumably it will also include standard measurement models, statistical techniques and so on. In addition, Alexandrova emphasizes the inclusion of mid-level theories, which she hopes will do the job of justifying scientists' uses of well-being.

The use of mid-level theories in well-being is not new (e.g. Giuntoli *et al*. 2015); they are also very similar to what measurement psychologists call construct theories (Stenner *et al*. 1983). Alexandrova does not draw attention to previous deployments of mid-level theories in well-being. This omission obscures her contribution, which, in my view, is to place some of the responsibility for developing these theories at the feet of philosophers. For too long there has been a chasm between the abstract high theorizing of philosophers, and the empirical data and expert opinions of well-being scientists. And although, increasingly, measurement psychologists recognize the importance of construct theories, with the exception of reading and maths ability, these kind of theories are slow to develop. Moreover, discussions of construct theories rarely include recognition of the normativity of well-being constructs.

For Alexandrova, mid-level theories in philosophy should aim to move among scientific fact, expert opinion and abstract philosophical theories. She uses the example of child well-being. She begins by taking account of expert beliefs about childhood well-being with the eye to using these beliefs as constraints on philosophical theories. Experts believe that objective indicators are important for measuring child well-being and that the concept of development is key to understanding childhood. Alexandrova uses these and other beliefs to test different philosophical

theories: ideal subjectivism is set aside largely due to the importance of objective indicators; Richard Kraut's developmentalism, a type of objective list theory, however, is pursued in part because it makes room for the development experts understand as important to childhood well-being.

In moving from these considerations to a useable theory, Alexandrova and her colleague Ramesh Raghavan provide the following mid-level theory of childhood well-being:

(1) Develop those stage-appropriate capacities that would, for all we know, equip them for successful future, given their environment;
(2) Engage with the world in child-appropriate ways, for instance, with curiosity and exploration, spontaneity, and emotional security.

As Alexandrova freely admits, this theory may still be too 'high', i.e. not sufficiently applied for some research contexts, but it is a good beginning.

In developing mid-level theories, Alexandrova argues that philosophers need to be prepared to stray from the tradition of including only non-instrumental goods in their theories. For mid-level theories to be useable and to justify the concepts scientists employ, they need to be specific. With that specificity may come instrumental goods. With the mixing of instrumental and non-instrumental goods, Alexandrova notes that we also tend to mix facts and values. For instance, when we begin to think about childhood and the empirical facts that attend to it – for example that it is a time of development – these facts in part tell us what is good for children – for example children should be safe and free from poverty. Alexandrova argues that the science of well-being is laden with values in a particular way that is different from other ways that values have been discussed in philosophy. She calls these value-laden claims mixed claims. Specifically, a claim is mixed if and only if:

(1) It is an empirical claim about a putative causal or statistical relation; and
(2) At least one of the variables in this claim is defined in a way that presupposes a moral, prudential or political value judgement about the nature of this variable.

Those familiar with well-being research will recognize these 'mixed claims' as the kind of claim that almost every empirical study makes – they are typically, although not exclusively, the conclusions of a study. Alexandrova provides pithy examples such as 'Happiness is not always conducive to well-being', but they are not always this obvious (e.g. Jeevan *et al*. 2017).

Alexandrova spends a lot of time defending the uniqueness of mixed claims as a category in philosophy and their legitimate place

in well-being science. Whether or not mixed claims are unique,[1] their legitimacy is at least partly a function of their widespread use in science. Thus, to my mind, Alexandrova's most important point is to note the danger of scientists' inattention to the normative nature of well-being concepts. Unacknowledged values can shape well-being science in ways that consumers of the science may find problematic. She argues, and I agree, that to be objective, i.e. to give the community reasons to trust this science, well-being scientists need to be more explicit about the values that play a role in concept and construct formation. Moreover, she argues that scientists should ask whether their concepts of well-being are robust in the face of different theories. If not, then the measuring instruments embodying these constructs should face public scrutiny and deliberation.

In the final two chapters Alexandrova turns to issues in measurement. Measurement is a key tool in the science of well-being, indeed, measurement is arguably the primary product of well-being science. But many have questioned the measurability of well-being. Can we quantify such a concept? Measurability has a long history in measurement theory, which unfortunately Alexandrova does not discuss. The central question is: what kind of evidence is needed to establish mappings between numbers and that which one is attempting to measure (see Tal 2017).

Alexandrova's main interlocutor in this chapter is Dan Hausman (2015) who does not believe well-being is measurable. Hausman commits to a conception of measurement that requires at least an interval scale. To establish the measurability of well-being on such a scale, we need evidence that it can be constructed as a quantitative variable, i.e. the difference between its values is equal. Hausman's argument against measurability relies heavily on the heterogeneity inherent in well-being: what is good for an individual is dependent on her values, identity, etc. Individuals clearly have different values, identities and so on. Thus, Hausman might say that well-being, when considered as a concept applicable to populations, is too heterogeneous to be construed as a quantitative variable.

Arguing against Hausman, Alexandrova suggests that even if general well-being is not measurable, perhaps context-specific well-being is, for example upper limb functioning. Perhaps, if we limit ourselves to the goods that apply to a particular kind of people or particular site of well-being, it will be easier to find legitimate and stable contributions to it. Indeed, this is what mid-level theories are supposed to supply. Curiously, she does not discuss the specific kind of evidence one would need to

[1] In their paper 'The Geography of Epistemic Risk' Justin Biddle and Rebecca Kukla (2017) discuss what they call phonetic risk, a category of epistemic risk that may include the heart of Alexandrova's 'mixed claims'.

https://doi.org/10.1017/S026626711800055X Published online by Cambridge University Press

ensure that the concept involved in the mid-level theory is measurable – a discussion that takes up significant space in the well-being literature amongst those who value interval level measurement (Michell 1999; Cano and Hobart 2011). Instead, Alexandrova moves directly to emphasize the importance of construct validation. This move seems too quick. Interval level measurement – if this is what Alexandrova is after – is difficult to achieve in well-being, and it is not clear that construct validation, even the robust validation she advocates, can serve to meet its conditions.

For Alexandrova, measurement in well-being science is possible when we combine a context-specific construct of well-being with construct validation. To be construct valid on Alexandrova's account, a measure must respect three types of evidence, what she calls the implicit logic of construct validation:

(1) It must be inspired by a plausible theory of its construct. This theory should be articulated as fully as possible and defended against alternatives;
(2) Subjects' responses to the questions in the measure reveal that they track the construct;
(3) Other knowledge of the construct is consistent with variations in values of the measure across contexts. This knowledge should encompass normative significance of the construct including moral and political contexts of its use.

As Alexandrova acknowledges, the problem with this logic is that in practice, measures of well-being rarely meet these conditions. In particular, construct validation tends to avoid condition 1 and the second part of condition 3. The psychometric tests used by psychologists to ensure the validity of their measures give lip service to theory, but in practice avoid it. Alexandrova blames what she calls 'evidentiary subjectivism' for this problem. She defines evidentiary subjectivism as:

> φ should be accepted as a component of x's well-being only if φ is a self-report of a factor shown to capture the data of well-being questionnaires completed by subjects relevantly similar to x (behavioural subjectivism) *or* φ is a self-report of a factor systemically claimed to be valued by subjects relevantly similar to x in a well-designed interview or survey (conversational subjectivism). (141)

On Alexandrova's account, well-being scientists overly rely on representative samples of the general population to determine the content of a measure, i.e. content validity. Moreover, in the context of psychometric tests, such as factor analysis, these tests rely on respondent answers to questions to determine what is or is not part of the concept of well-being under investigation. The problem with evidentiary subjectivism, according to Alexandrova, is that it avoids, and at the same time replaces,

https://doi.org/10.1017/S026626711800055X Published online by Cambridge University Press

normative theorizing. This is particularly problematic as evidentiary subjectivism does not have a mechanism for correcting when respondents reason wrongly about well-being, nor is there a way to ensure that respondent answers reflect what they truly value. Alexandrova is very critical of evidentiary subjectivism, going so far as to say that there is no justification for it. While I agree with some of her insights, I think the use of first-person reports in well-being measurement is more complex than this book admits. Sometimes they are an attempt at normativity. But without detailed examples, including how these measures are used, it is difficult to assess this complexity.

Alexandrova suggests that to overcome the problem of evidentiary subjectivism, scientists need to engage with normative theory in both conditions 1 and 3 of her implicit logic. On the one hand, this suggestion is realistic in the sense that well-being scientists are not entirely unaware nor uncritical of their lack of theorizing (Cano & Hobart 2011). On the other hand, progress is slow and Alexandrova is right not to underestimate the difficulty of getting scientists to engage in normative theory.

Leah McClimans*

REFERENCES

Biddle, J. B. and R. Kukla 2017. The geography of epistemic risk. In *Exploring Risk: Case Studies of Values in Science*, ed. K. E. Elliott and T. Richards, 215–38. Oxford: Oxford University Press.

Cano, S. J. and J. C. Hobart 2011. The problem with health measurement. *Patient Preference and Adherence* 5: 279–290.

Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Cartwright, N., T. Shomar and M. Suarez. 1995. The tool box of science: tools for the building of models with a superconductivity example. In *Theories and Models in Scientific Processes*, ed. W. E. Herfel, W. Krajewski, I. Niiniluoto and R. Wojcicki, 137–149. Amsterdam: Rodopi.

Giuntoli, G., S. Huges, K. Karban and J. South. 2015. Towards a middle-range theory of mental health and well-being effects of employment transitions: findings from a qualitative study on unemployment during the 2009–2010 economic recession. *Health* 19: 389–412.

Hausman, D. 2015. *Valuing Health*. Oxford: Oxford University Press.

Jeevan, R., J. P. Browne, C. Gulliver-Clarke, J. Pereira, C. M. Caddy, J. H. P. van der Meulen and D. A. Cromwell. 2017. Surgical determinants of patient-reported outcomes following postmastectomy reconstruction in women with breast cancer. *American Society of Plastic Surgeons* 139: 1036e–1045e.

Michell, J. 1999. *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge: Cambridge University Press.

* University of South Carolina and University College Cork, College Rd, University College, Cork T12 K8AF, Ireland. Email: lmm@sc.edu. URL: https://sc.edu/study/colleges_schools/artsandsciences/our-people/faculty-staff/mcclimans_leah.php.

https://doi.org/10.1017/S026626711800055X Published online by Cambridge University Press

Stenner, J., M. Smith III and D. S. Burdick. 1983. Toward a theory of construct definition. *Journal of Educational Measurement* 4: 305–316.
Tal, E. 2017. Measurement in science. In *Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. <https://plato.stanford.edu/archives/fall2017/entries/measurement-science/>.

BIOGRAPHICAL INFORMATION

Leah McClimans is an Associate Professor of Philosophy at the University of South Carolina and an ASSISTID Marie Curie Fellow in the School of Nursing and Midwifery at University College Cork. She is the editor of *Measurement in Medicine: Philosophical Essays on Assessment and Evaluation* (Rowman and Littlefield). She has written extensively on quality of life measures in health care and values in evidence-based medicine.

doi:10.1017/S0266267119000051

The Philosophy of Social Evolution, Jonathan Birch. Oxford University Press, 2018, xi + 268 pages.

The Philosophy of Social Evolution is an engaging investigation of the intellectual legacy of Bill Hamilton, a central figure in evolutionary biology whose ideas have had a lasting impact far beyond that discipline. The book combines mathematical rigour and accessibility, as many of the more mathematical points can be skipped without too much loss. Birch provides a well-reasoned and even-handed discussion of some controversial topics surrounding social evolution in the first part of the book. Then, he offers suggestions as to how these ideas may be extended and used in areas beyond where they have traditionally been applied, such as public goods games in both microbial and human populations. While this book is foremost a work in philosophy of biology, it should also be of interest to broader audiences, including philosophers of economics, as many of the foundational concepts Birch discusses have found their way into other fields (philosophy, anthropology, economics, sociology, etc.) and some of the extensions proposed (especially those in Chapter 8) are relevant to studying human behaviour.

Chapter 1 provides some examples of social traits, then explains Hamilton's four-part schema which categorizes social behaviours (dividing them into mutually beneficial, selfish, altruistic and spiteful) based on their effects on the reproductive success of the organism and its social partner(s). While the schema is nothing new, Birch argues that the best way to understand the categories is in terms of a trait's recent selective history (e.g. it is an altruistic trait if it was recently selected for its benefits to an organism's social partners, despite having a cost for the altruistic organism itself). He also argues that we should