Measurement Theory and the Foundations of Utilitarianism

John A. Weymark
Department of Economics, Box 1819, Station B,
Vanderbilt University, Nashville, TN 37235, U.S.A.
(e-mail: john.weymark@vanderbilt.edu)

Extended Abstract. Harsanyi [“Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility”, *Journal of Political Economy*, 1955] used expected utility theory to provide two axiomatizations of weighted utilitarian rules. Sen [“Welfare inequalities and Rawlsian axiomatics”, *Theory and Decision*, 1976] has argued that Harsanyi has not, in fact, axiomatized utilitarianism because he has misapplied expected utility theory. Specifically, Sen has argued that von Neumann–Morgenstern expected utility theory is an ordinal theory and, therefore, any increasing transform of a von Neumann–Morgenstern utility function is a satisfactory representation of a preference relation over lotteries satisfying the expected utility axioms. In other words, von Neumann–Morgenstern expected utility is ordinal. However, Harsanyi's version of utilitarianism requires a cardinal theory of utility in which only von Neumann–Morgenstern utility functions are acceptable representations of preferences. Sen's informal discussion of these issues has been formalized by Weymark [“A Reconsideration of the Harsanyi–Sen debate on utilitarianism', in Elster and Roemer, eds., *Interpersonal Comparisons of Well-Being*, 1991].

Broome [*Weighing Goods*, 1991, and “Can there be a preference-based utilitarianism?”, forthcoming in Salles and Weymark, eds., *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls*] has argued that von Neumann-Morgenstern expected utility theory is cardinal in the relevant sense needed to support Harsanyi's utilitarian conclusions. His basic point is that a preference binary relation is not a complete description of preferences in the von Neumann-Morgenstern theory. Rather, the preference relation needs to be supplemented by a binary operation, and it is this operation that makes the theory cardinal.

Broome does not provide a formal argument in support of this conclusion. Instead, he argues by analogy with how measurement theory constructs scales for weighing objects (and for measuring temperature). To construct

a measure of weight, one starts with a set of objects S, a binary relation $\succeq$ on S (weighs at least as much as), and a concatentation operation $\circ$ from $S \times S$ to S (the objects x and y are combined to create a new object $x \circ y$, e.g., by placing them in the same pan of a balancing scale). A real-valued function $w\colon S \to \mathbb{R}$ is a weight measurement scale if w both represents $\succeq$ and $w(x \circ y) = w(x) + w(y)$.

I use formal measurement theory to critically evaluate Broome's argument. In measurement theory, a *relational structure* is a set S together with one or more relations on S. (Binary operators on $S \times S$ can be equivalently thought of as ternary relations on S.) Constructing a measurement scale amounts to establishing a homomorphism between an empirical relational structure and a numerical one.

For example, the empirical relational structure used to measure weight is $\mathcal{W} = \langle S; \succeq; \circ \rangle$, where S, $\succeq$, and $\circ$ are as defined above. The usual practice in the theory of weight is to associate the numerical structure $\mathcal{N} = \langle \mathbb{R}; \geq; + \rangle$ with $\mathcal{W}$. The homormorphism is a function f that maps S into $\mathbb{R}$ in such a way that the properties of $\succeq$ and $\circ$ are preserved by $\geq$ and $+$. Such a function is unique up to a similarity transform, from which it is often asserted that weight is a cardinal concept. However, as noted by Krantz, Luce, Suppes, and Tversky [*Foundations of Measurement, Volume I*, 1971] the choice of an additive representation in this case is simply a matter of convenience. For example, we could instead use $\bar{\mathcal{N}} = \langle \mathbb{R}; \geq; \times \rangle$ as the numerical structure. If this is done, then we would want to have $f(x \circ y) = f(x)f(y)$, rather than $f(x \circ y) = f(x) + f(y)$, as would be the case with $\mathcal{N}$.

In von Neumann-Morgenstern expected utility theory, the empirical relational structure is $\mathcal{E} = \langle \mathcal{L}; \succeq; \alpha, \alpha \in (0, 1) \rangle$, where $\mathcal{L}$ is a set of lotteries on a set of alternatives, $\succeq$ is a binary relation on $\mathcal{L}$, and each α is a binary operator on $\mathcal{L} \times \mathcal{L}$ that maps the ordered pair of lotteries (p, q) to the lottery $\alpha p + (1 - \alpha q)$. [Strictly speaking, von Neumann and Morgenstern work with a set of 'abstract utilities' $\mathcal{U}$, rather than a set of lotteries. I follow the subsequent literature in describing their theory in terms of lotteries.]

Broome never makes clear that $\mathcal{E}$ is the empirical relational structure for expected utility theory corresponding to $\mathcal{W}$ for weight. Nor does he explain why measuring weight is analogous to measuring utility in risky situations. Given that the operators $\circ$ and α are different, it is not obvious that the analogy is a good one.

Broome's argument implicitly assumes that the numerical relational structure corresponding to $\mathcal{E}$ is $\mathcal{V} = \langle \mathbb{R}; \geq; \alpha^*, \alpha^* \in (0, 1) \rangle$, where α^* is the convex operator that maps the pair of numbers (a, b) into $\alpha^* a + (1 - \alpha^*) b$. With this numerical relational structure, in order for α^* to preserve the structure of α, a representation of $\succeq$ would be acceptable if and only if it is a von Neumann-Morgenstern utility function. Such a representation is unique up to an increasing affine tranform, from which it might be claimed that von Neumann-Morgenstern utility theory is cardinal in the sense required by utilitarianism.

However, as with the measurement of weight, the choice of the numerical relational structure in the preceding discussion has been chosen because the operators that appear in the numerical relational structure appear to be 'natural' analogues of the operators that appear in the empirical relational structure. This is perfectly acceptable in positive theories in which one is merely trying to work out the implications of expected utility theory for choice behaviour. However, it is not a satisfactory basis for choosing a numerical relational structure when the objective is to provide a normative foundation for utilitarianism. What is missing is an argument that the von Neumann-Morgenstern utility representations of $\succeq$ obtained using the numerical relational structure $\mathcal{V}$ (rather than the representations obtained using binary operators different from the α^*) are the ethically-relevant ones for making the interpersonal comparisons of utility differences required by utilitariansm. It is simply not possible to do this without enriching the theory in some way that goes beyond what is in the von Neumann-Morgenstern formulation of expected utility theory.

These observations show that Sen's initial criticism of Harsanyi can be extended to the more complete description of expected utility theory that is obtained by using the mixture operators α in addition to the preference relation $\succeq$.

In the full version of this article, these arguments will be spelled out more formally and in greater detail. Further examples taken from the measurement theory literature will be used to illustrate the arguments (e.g., how probabilities are measured in Savage's theory). There will also be discussion of versions of Broome's argument that have appeared in the work of Binmore [*Game Theory and the Social Contract: Volume 1*, 1994] and Risse ["Harsanyi's 'Utilitarian Theorem' and Utilitarianism", *Noûs*, 2002]. An alternative approach to cardinal utility that uses degrees of preference (which has been advanced by Harsanyi in unpublished correspondence) will also be considered.