

Is it possible to measure happiness?

The argument from measurability

Erik Angner

Received: 13 August 2012 / Accepted: 3 March 2013 / Published online: 5 April 2013
© Springer Science+Business Media Dordrecht 2013

Abstract A ubiquitous argument against mental-state accounts of well-being is based on the notion that mental states like happiness and satisfaction simply cannot be measured. The purpose of this paper is to articulate and to assess this “argument from measurability.” My main thesis is that the argument fails: on the most charitable interpretation, it relies on the false proposition that measurement requires the existence of an observable ordering satisfying conditions like transitivity. The failure of the argument from measurability, however, does not translate into a defense of mental-state accounts as accounts of well-being or of measures of happiness and satisfaction as measures of well-being. Indeed, I argue, the ubiquity of the argument from measurability may have obscured other, very real problems associated with mental-state accounts of well-being – above all, that happiness and satisfaction fail to track well-being – and with measures of happiness and satisfaction – above all, the tendency toward *reification*. I conclude that the central problem associated with the measurement of, e.g., happiness as a subjectively experienced mental state is not that it is *too hard* to measure, but rather that it is *too easy* to measure.

Keywords Well-being · Welfare · Happiness · Measurement · Reification

1 Introduction

After a spectacular run lasting several hundred years, the twentieth century saw the precipitous decline of mental-state accounts of well-being – accounts according to which a person is well off to the extent that she is in some particular, subjectively experienced mental state, like happiness or satisfaction (Sumner 1996, p. 83). Instead, philosophers by and large came to favor preference-satisfaction accounts – according to which a person is well off to the extent that her preferences are satisfied

E. Angner (✉)

Department of Philosophy, George Mason University, 4400 University Drive 3F1,
Fairfax, VA 22030, USA
e-mail: eangner@gmu.edu

(Rabinowicz and Österberg 1996) – or objective-list accounts – according to which certain things are good or bad for people largely independently of whether those people want those things or whether they would give rise to positive or desirable mental states (Nussbaum 2008; Sen 1985). A ubiquitous argument against mental-state accounts is based on the notion that subjectively experienced mental states like happiness and satisfaction simply cannot be measured. As Tim Chappell and Roger Crisp write in the *Routledge encyclopedia of philosophy*: “Desire theories have come to dominate contemporary thought because of economists’ liking for the notion of ‘revealed preferences’ ... Pleasures and pains are hard to get at or measure, whereas people’s preferences can be stated, and inferred objectively from their behaviour” (Chappell and Crisp 1998, p. 553).

As this quote suggests, the shift away from mental-state accounts of well-being parallels, and may in fact have been driven by, developments in economics. Until the beginning of the twentieth century, economists largely thought about welfare or well-being – I use the terms interchangeably – as a matter of subjectively experienced mental states. As A. C. Pigou put it in *The economics of welfare*, originally published in 1920: “[The] elements of welfare are states of consciousness and, perhaps, their relations” (Pigou 1952, p. 10). During the first half of the twentieth century, however, economists by and large abandoned the view that welfare had any essential connection with subjectively experienced mental states. Under the influence of logical empiricism in philosophy, behaviorism in psychology, and operationalism in physics, and unimpressed with the state of their own discipline, economists came to think of references to unobservable mental states like happiness and satisfaction as unscientific and at any rate dispensable (Angner and Loewenstein 2012; Mandler 1999). Hence, mainstream economists abandoned mental-state accounts of well-being in favor of preference-satisfaction accounts (implicit in all the orthodox economic welfare measures, including consumer and producer surplus, as well as compensating and equivalent variation) or objective-list accounts (implicit, e.g., in the Human Development Index of the United Nations Development Programme) (Angner 2009a).

Arguments like that of Chappell and Crisp appear in the writings of both philosophers and economists. Philosophers Christoph Fehige and Ulla Wessels, like Chappell and Crisp, present the case as an argument for the superiority of preference-satisfaction over mental-state accounts of well-being:

Yet another argument for preferences is that from *comparability and measurability*. A welfarist will, in some sense or other, want to compare or quantify welfare. “Causing people great pain is worse than causing them mild pain”, is one of the things she will want to say. The most promising road to comparability seems to lead, even in the case of pleasures and pains, via the agent’s preferences: “Your possible headache *X* is milder than your possible toothache *Y*”, we could say, “if, other things being equal, you would prefer having *X* to having *Y*” (Fehige and Wessels 1998, p. xxxvi).

Clearly, this “argument for preferences” proceeds from the notion that degrees of preference satisfaction are measurable, whereas degrees of happiness are not. Similarly, in his defense of Gross National Product (GNP) as a welfare measure, economist Wilfred Beckerman (1975) recognizes that there are different conceptions of welfare and acknowledges that some people think welfare is a matter of happiness.

Nevertheless, he objects: “[No] science can tell us whether modern man is happier than mankind a hundred years ago, or even 10 years ago. The concept of happiness is one for which there can be no scientific objective measure” (Beckerman 1975, p. 53). Implicit in these arguments is the premiss that an adequate account of well-being must entail that well-being is measurable. This proposition, which I will call the *measurability requirement*, is quite widely shared. As Thomas Scanlon describes it:

Well-being is ... commonly supposed to be a notion that admits of quantitative comparisons of at least some of the following kinds: comparisons of the levels of well-being enjoyed by different individuals under various circumstances, comparisons of the increments in a single individual’s well-being that would result from various changes, and perhaps also comparisons of the amounts of well-being represented by different lives, considered as a whole (Scanlon 2000, p. 109).

The measurability requirement is in part motivated by ethical theories like utilitarianism, which require a notion of well-being that is both measurable and interpersonally comparable.

Meanwhile, increasing numbers of social and behavioral scientists reject the notion that happiness and satisfaction, understood as subjectively experienced mental states, are unmeasurable (cf. Brooks 2008, p. 9; Gilbert 2006, p. 64). Bruno S. Frey and Alois Stutzer insist that happiness is measurable not just in principle, but in practice:

Recently, great progress has been achieved in economics: happiness has been seriously measured, and many of its determinants have been identified. This constitutes a sharp break from the notion, much cherished by economists, that revealed preferences only reliably reflect individual utility, and that it is the only way to make serious measurements (Frey and Stutzer 2000, p. 145).

Under headings like “positive psychology” and “the science of happiness,” researchers like Frey and Stutzer explore the determinants and distribution of happiness, satisfaction, and other “positive” or desirable subjectively experienced mental states. Recognizing that the enterprise depends critically on the notion that such states can be validly and reliably measured, social and behavioral scientists have developed a variety of measures (Angner 2009a, b). Since the term “subjective well-being” has come to refer to whatever positive or desirable state attracts an author’s attention, the measures are frequently called “measures of subjective well-being”; since such authors often equate well-being and subjective well-being, the measures are also called “subjective measures of well-being” (Angner 2010). As measures of well-being, subjective measures are frequently presented as substitutes for, or complements to, orthodox economic measures for public policy purposes (Diener 2000, 2006; Diener and Seligman 2004; Diener et al. 2009; Kahneman et al. 2004a, b). The contrast between these social and behavioral scientists, on the one hand, and critics of mental-state accounts, on the other, suggests that a reassessment of the argument from measurability is overdue. Since the reassessment will hinge on assumptions about what can and cannot be measured, it is a job (in part) for the philosophy of science.

Why does the argument from measurability matter? Obviously enough, the argument from measurability constitutes a major threat to mental-state accounts of well-being: if the argument is sound, mental-state accounts of well-being are inadequate. Mental-state accounts are having a real renaissance in the philosophical literature, as a number of

philosophers have been inspired by the empirical literature to include at least some subjective elements into their accounts of well-being (e.g., Haybron 2008; Sumner 1996; Tiberius and Plakias 2010), and all of these accounts are potentially vulnerable to the argument from measurability. The argument also represents a serious challenge to the science of happiness: if happiness and other mental states cannot be measured, the effort to understand scientifically the distribution and determinants of happiness would presumably be in vain. Finally, the argument would also seem to undercut the effort to reorient public policy in light of the science of happiness: if mental-state accounts are inadequate, this would be a challenge to the notion that subjective measures represent well-being and tell us anything about the welfare implications of policy interventions. Because subjective measures suggest rather different answers to questions about the determinants and distribution of well-being as compared to economic indicators, the choice has considerable policy implications.

The aim of the present paper is to articulate and assess what Fehige and Wessels call the argument from measurability.¹ My main thesis is that, on the most charitable interpretation, the argument from measurability fails because it relies on a false premiss, viz., the proposition that measurement requires the existence of an observable ordering that satisfies conditions like transitivity. As it turns out, the proposition is virtually unanimously rejected in contemporary social and behavioral science and for good reason. The failure of the argument from measurability, however, does not translate into a defense of mental-state accounts as accounts of well-being or of measures of happiness and satisfaction as measures of well-being. Indeed, I argue, the ubiquity of the argument from measurability may have obscured other, very real problems associated with mental-state accounts of well-being – above all, that happiness and satisfaction fail to track well-being – and with measures of happiness and satisfaction – above all, the tendency toward *reification*. I conclude that the central problem associated with the measurement of, e.g., happiness as a subjectively experienced mental state is not that it is *too hard* to measure, but rather that it is *too easy* to measure.

2 The measurement of happiness and satisfaction

What I have called “subjective measures” includes a range of specific measures (Angner 2011a). For most of their history, subjective measures were constructed on the basis of one or more direct questions like: “Taking things all together, how would you say things are these days – would you say you’re *very happy*, *pretty happy*, or *not too happy* these days?” (Gurin et al. 1960, p. 411). Sonja Lyubomirsky and Heidi S. Lepper (1999) offer four prompts, including “In general, I consider myself,” and invite subjects to respond on a seven-point scale where 1 represents “In general, I consider myself not a very happy person” and 7 “In general, I consider myself a very happy person” (Lyubomirsky and

¹ By contrast, I will ignore what Fehige and Wessels call “the argument from comparability” and I will not challenge the measurability requirement. Not everyone agrees with it: Scanlon (2000) distances himself from the view outlined in the quote above, and Amartya Sen (e.g., 2002) has argued that incomplete orderings need not be an obstacle to social choice. Yet, by and large, proponents of subjective measures of well-being appear to endorse it: “science is about measurement, and if a thing cannot be measured – cannot be compared with a clock or a ruler or something other than itself – it is not a potential object of scientific inquiry” (Gilbert 2006, p. 64).

Lepper 1999, p. 151). Others ask subjects “How do you feel about your life as a whole?” and give them response categories ranging from “Delighted,” “Pleased,” and “Mostly satisfied,” through “Mixed (about equally satisfied and dissatisfied)” to “Mostly dissatisfied,” “Unhappy” and “Terrible” (Andrews and Withey 1976, p. 18). Occasionally, researchers have elicited responses using graphic representations like horizontal lines (Watson 1930), ladders and mountains (Cantril 1965), or happy and sad faces (Andrews and Withey 1976). Some of these questions were designed to represent affective states, some to represent cognitive states, and some to represent some combination of affective and cognitive states (Angner 2010).

A somewhat different approach has been developed by Daniel Kahneman (1999) and others under the heading of *experience sampling*. Kahneman prompts his subjects every so often – e.g., with the use of handheld devices – to judge the “quality of their momentary experience” along the “good/bad dimension” (Kahneman 1999, p. 7). The assumption is that, at every point in time, the brain rates the quality of experience in a manner that can be represented on a single numerical scale and which, furthermore, is accessible to the agent. What matters, at the end of the day, is the time integral (which Kahneman calls *objective happiness*) of instant happiness ratings (which Kahneman calls *subjective happiness*) (Kahneman 1999, p. 5). The effort to produce a dense record of an individual’s affective state as a function of time was pioneered by Hornell Hart (1940), the inventor of the *Euphorimeter*: a device that would permit the quick assessment of an individual’s level of happiness based on self-reports.

More recently, Kahneman and Alan B. Krueger have suggested the use of a measure they call the *U-index*, which they present as a measure of society’s well-being (Kahneman and Krueger 2006). The “U” stands for “unpleasant” or “undesirable,” and the index “measures the proportion of time an individual spends in an unpleasant state,” where an episode gets classified as pleasant or unpleasant depending on whether the strongest affect experienced during the episode is positive or negative (Kahneman and Krueger 2006, pp. 18–19). The U-index was designed to overcome several perceived problems associated with other subjective measures, above all problems related to interpersonal comparability (Krueger 2009, p. 3). Still, Kahneman and co-authors insist: “Experience sampling is the gold standard” (Kahneman et al. 2004a, p. 1777).

Why would anyone think that subjectively experienced mental states like happiness and satisfaction can be measured? Key to answering this question is to understand that proponents of subjective measures operate with the so-called *psychometric* approach to measurement. This approach emphasizes latent constructs and construct validity (Angner 2012). A latent variable, or *construct*, according to a standard textbook, is “a variable [that] is abstract and latent rather than concrete and observable” (Nunnally and Bernstein 1994, p. 85). On the psychometric approach, you start off by simultaneously postulating the existence of a construct and proposing a measure of it. Then, you explore patterns of variances and covariances: relationships between the proposed measure and (a) measures of (i) other constructs and (ii) overt behavior, as well as (b) across conditions. If the pattern of variances and covariances conforms to expectations – that is, if the measure “behaves as expected” – you infer that the construct has been validated; if not, you infer either that there is something wrong with the underlying construct or with the measure itself, and start over. The process, referred to as *construct validation*, is often described as an instance of hypothesis testing (Cronbach and Meehl 1955, p. 300;

Johnson 2001, p. 11316). In brief, when defending a given measure, those operating in accordance with the psychometric approach reason inductively from the claim that a given measure behaves as expected.

The defense of measures of happiness and satisfaction conforms perfectly to this picture. Proponents of such measures postulate the existence of a construct – happiness, satisfaction, or similar – and propose a measure of it, and then examine patterns of variances and covariances to determine whether the measure “behaves as expected” when compared to objective life circumstances, other people’s judgment of subjects’ happiness, measures of mental health, and so on (cf. Lyubomirsky and Lepper 1999, p. 145). In an early review of the literature, Warner Wilson made the case by arguing that self-reported happiness scores were sufficiently correlated with the judgments of associates, teachers, professors, principals, psychologists, and clinical judges, as well as with average scores on elation-depression scales (Wilson 1967, pp. 294–295). Hart found that happiness scores, tracked over time, changed as expected when his participants fell in love, experienced the death of their mothers, or contemplated suicide (Hart 1940, pp. 19–25). According to a more recent review, subjective measures are sufficiently positively correlated with happiness ratings of friends and family, psychologists’ judgments, amount of smiling; sufficiently negatively correlated with depression; and not overly correlated with general intelligence, current mood, humility, and the language in which the question was asked (Diener and Suh 1997, pp. 436–438). The fact that the defense of subjective measures satisfies the strictures imposed by the psychometric approach is not coincidental: psychometrics and happiness measurements can both trace their historical roots to the emergence of personality psychology during the aftermath of WWI (Angner 2011a).

The important thing to notice is that within the psychometric approach, there is no principled reason to think that unobservable mental states like happiness and satisfaction cannot be measured. In fact, the psychometric approach was developed precisely to quantify unobservable individual differences in intelligence and other forms of “mental ability,” “mental functioning,” etc. (Jones and Thissen 2006, pp. 5–6). As we have seen, psychometrics does allow for the possibility that measurement might fail: if a proposed measure does not behave as expected, its author would be required to infer either that there is something wrong with the underlying construct or with the measure itself. Yet, within this approach, the possibility that an unobservable entity like a subjectively experienced mental state is measurable cannot be summarily dismissed, as those who endorse the argument from measurability believe. We can be quite confident, then, that those who endorse the argument from measurability do not operate with the psychometric approach to measurement.

3 The argument from measurability

As it happens, there are two different approaches to measurement in social and behavioral science (Dawes and Smith 1985; John and Benet-Martínez 2000; Judd and McClelland 1998; Krantz 1991). As David H. Krantz puts it: “One, which may be termed the psychometric approach, introduces latent variables to explain behavioral orderings. The second ... treats the numerical representation of behavioral orderings axiomatically” (Krantz 1991, p. 2). The second approach, which is called the

measurement-theoretic or *representational* approach, emphasizes observable orderings, homomorphisms, and representation theorems (Angner 2012). As we will see, the differences between the psychometric and the measurement-theoretic approaches have deep implications for the nature of measurement in general and for the measurability of subjectively experienced mental states in particular.²

On the measurement-theoretic approach, you start off with a set A of objects (e.g., rods, commodity bundles), which can be ordered with respect to some property (e.g., length, preference) by applying a simple observable operation. This observable ordering will as a matter of fact satisfy certain conditions. These conditions, which can be established by applying the observable operation, can be expressed as set of axioms (Krantz et al. 1971, p. 6), which can be thought of as a set of empirical laws (Krantz et al. 1971, p. 13). Then, you offer a representation theorem: a proof to the effect that if the empirical relation \succ satisfies certain properties, then there is a function $\phi(\cdot)$ from A into some set of numbers such that $\phi(\cdot)$ is a *homomorphism*, that is, an assignment of numbers to each member of A such that one object bears relation \succ to another just in case the former is associated with a greater number than the latter (Krantz et al. 1971, p. 9). In brief, those operating in accordance with the measurement-theoretic approach reason deductively from the claim that a given empirical relation satisfies certain axioms.

The defense of orthodox economic welfare measures conforms perfectly to this picture. Orthodox economists begin with the assumption that observed choices over some set of alternatives satisfy certain axioms. On the basis of the axioms, they offer a formal proof (called a representation theorem) to the effect that the measure (called a utility function) is a homomorphism. This way, “utility” is divorced from any association with subjectively experienced mental states: in technical terms, a utility function is simply an *index*, or *measure*, of preference satisfaction. Proofs are available in any standard-issue graduate-level microeconomics textbook like Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green’s *Microeconomic theory* (1995). The fact that the defense of orthodox economic measures satisfies the strictures imposed by the measurement-theoretic approach is not coincidental: the measurement-theoretic approach was in fact developed in large part to solve the problem of utility measurement (Krantz et al. 1971, p. 9).

The important thing to notice is that the measurement-theoretic approach to measurement, given certain common assumptions, entails that degrees of preference satisfaction are measurable, whereas degrees of happiness are not. According to the measurement-theoretic approach, measurement is possible if and only if there exists a suitable observable ordering satisfying conditions like transitivity. As Norman Cliff puts it: “Measurement theory says that if certain conditions hold, then scales of a given kind are defined. If not, they are not” (Cliff 1992, p. 189). In the case of utility measurement, the ordering in question is that imposed on a set of options by people’s choices, and the conditions are the axioms of rational choice (Krantz 1991, p. 28). While the axioms are frequently treated as normative laws of rational choice, the point here is that for the representation theorem to get traction, they must be true descriptive laws. Given the assumption – central to orthodox economics, as evidenced by the treatment in Mas-

² See Angner (2012) for a more thorough discussion of the two approaches to measurement. The use of the term “theory of measurement” to refer to one specific approach to measurement might generate confusion, but it is the term favored in primary as well as secondary literature.

Colell et al. (1995) – that people are rational in the sense that their choices are consistent, observed choices constitute an ordering that can provide the foundation for utility measurement. On the additional assumption – which, to my knowledge, is shared by all – that there exists no such ordering in the case of happiness, degrees of happiness are unmeasurable.³ All of this suggests that those who endorse the argument from measurability operate with the measurement-theoretic approach.

My hypothesis, then, is that the argument from measurability is premised on the proposition, central to the measurement-theoretic approach, that measurement requires an observable ordering satisfying conditions like transitivity. The hypothesis accounts for the phenomena, since it explains why those who endorse the argument from measurability believe that degrees of preference satisfaction are measurable whereas degrees of happiness are not. Obviously, the hypothesis is underdetermined by the evidence: none of the authors cited in the introduction spells out a complete argument, so there are many logically possible ways to fill in the details. Yet, the fact that critics believe the measurability of subjectively experienced mental states can be dismissed so summarily is another phenomenon to be explained, and my hypothesis does: within the measurement-theoretic approach it is *obvious* that subjectively experienced mental states like happiness cannot be measured.

My hypothesis, moreover, is charitable. On my interpretation, the proposition that measurement requires observable orderings, far from an arbitrary supposition, is part and parcel of an approach to measurement that at the time was considered a revolutionary development and which more than half a century later continues implicitly to dominate mainstream economics. The approach has an excellent pedigree as it was worked out in great detail by top-notch scientists, logicians, and philosophers of science in venues like the *Journal of Symbolic Logic* (Scott and Suppes 1958) and in the three-volume *Foundations of measurement*.⁴ As Cliff puts it:

The people central to the development of abstract measurement theory are among the most creative and productive minds in scientific psychology, and abstract measurement theory has to be regarded as one of its major intellectual achievements, viewed in terms of the power of thought that was required to achieve it and as a key pillar of the philosophy of science (Cliff 1992, p. 186).

In addition, the measurement-theoretic approach to measurement, like neoclassical economics in general (see the [introduction](#)), was deeply influenced by a prevailing, broadly speaking empiricist philosophy of science (Trout 1998, p. 49). Because on my interpretation, the proposition that measurement requires observable orderings is no arbitrary supposition, but part and parcel of a prominent scientific development inspired by a prevailing philosophy of science, the attribution of this proposition to those who endorse the argument from measurability is both plausible and charitable.

There are other ways to spell out the argument from measurability. One possibility is to think of it as motivated by hostility to introspection of a kind common in twentieth-century science and philosophy. It is frequently argued that measures of

³ Not even the most ardent proponents of subjective measures assume the existence of such orderings. And for good reason: as we will see in section 6, empirical evidence suggests that happiness reports are too context dependent to serve the purpose.

⁴ The first volume appeared as Krantz et al. (1971).

affective states – including measures of happiness and the like – must rely on introspection (Alexandrova 2008, p. 572). Yet, recent work in philosophy of science has shown that much of the hostility is misguided (Piccinini 2009). And in this case, it is simply a misunderstanding to think that subjective measures of well-being necessarily rely on introspection: happiness measures rely on first-person reports, but they are silent on whether those reports are generated by introspective access to one's own mental states or, e.g., by Skinner-style observation of one's own behavior. Because this interpretation of the argument from measurability would attribute to critics of mental-state accounts a simple misunderstanding of the nature of happiness measures, this interpretation would be less charitable than the one outlined above. Even less charitable interpretations can easily be developed and dismissed; here, the goal is to assess the most charitable interpretation only.

It can be objected that those who endorse the argument from measurability could not possibly operate with the measurement-theoretic approach to measurement because they might be unaware of its existence. Similarly, it can be objected that my reading fails because no contemporary theorist would explicitly endorse the theory of measurement or the particular strain of empiricist philosophy of science underlying it. Yet, it is perfectly possible to operate with a given approach to measurement without being aware of it – that is, to operate with a given approach to measurement though not under that description – and to practice science within a set of methodological constraints without being able to name or even to identify them. By putting the argument from measurability in the context of a particular form of scientific practice as well as a particular strain of empiricist philosophy of science, we can understand where the argument came from, how it is to be spelled out, and why it exerted such pull.

4 Does measurement require observable orderings?

In the previous section, I maintained that the argument from measurability, on the most charitable interpretation, depends critically on the proposition that measurement requires the existence of an observable ordering satisfying conditions like transitivity. I take it for granted that this would make it impossible to measure subjectively experienced mental states, since not even the most ardent proponents of subjective measures maintain that such measures can be constructed on the basis of observable orderings satisfying conditions like transitivity. In this section, I will develop a two-pronged case against this proposition. First, I will argue that if the proposition were true, it would be an embarrassment not just for mental-state accounts of well-being, but for preference-satisfaction accounts as well, since real people's observable choices do not in fact constitute an observable ordering satisfying the relevant conditions either. Second, I will argue, even a cursory review of the best scientific practice strongly suggests that the proposition is false. Consequently, on the most charitable interpretation, the argument from measurability fails because it is unsound: it proceeds from the false premiss that measurement requires the existence of an observable ordering satisfying conditions like transitivity.

First, if measurement required an observable ordering satisfying conditions like transitivity, this would constitute a problem not just for mental-state accounts of well-being but for preference-satisfaction accounts as well. The best available empirical

evidence suggests that people's observable choices do not in fact satisfy such conditions. As Oliver P. John and Veronica Benet-Martínez put it:

Although representational measurement promised to provide a strong and defensible foundation for psychological measurement, it has so far failed to deliver on that promise. During the 1970s and 1980s a slew of studies, inspired by Tversky and Kahneman's pioneering work, showed that people's preferences, risk perceptions, political attitudes, and so on often violate the transitivity rule required ... and that judgments may shift substantially depending on the framing of the questions or items (John and Benet-Martínez 2000, p. 341).⁵

Cliff's article, titled "Abstract measurement theory and the revolution that never happened," argues that "formal measurement theory ... has so far had little influence on the rest of psychology" (Cliff 1992, p. 188). After noting that on this approach, measurement is possible if and only if observable orderings satisfy conditions like transitivity, he adds: "But data always contradict one or the other axiom" (Cliff 1992, p. 189). In a more recent handbook article on measurement, Charles M. Judd and Gary H. McClelland admit: "While there are some success stories in psychophysics for representational measurement, successful applications of representational measurement in social psychology are difficult to find" (Judd and McClelland 1998, p. 183). By 1991, Krantz himself had come to a very similar conclusion. Under the heading "The Myth of Utility," Krantz argues:

Choice does indeed depend on the method of testing and depends especially on how options are *framed*. Results such as these show that ordering options by choices is no more determinate than ordering "overall" reading skill by testing on a particular set of materials (Krantz 1991, p. 32).

Like John and Benet-Martínez, Krantz points to empirical results by Amos Tversky and Kahneman and their co-authors, whose work is widely interpreted as showing that real-life choices reflect what is often called "normatively irrelevant" factors and consequently fail to satisfy the relevant axioms. Krantz concludes: "Preference ordering is a behaviorist myth ... linked to the behavioral assumption that 'preferences' are 'revealed' by choices or 'elicited' by presentation of suitable options and to the mathematics of maximization" (Krantz 1991, p. 35).

From our vantage point, if anything, the case for the truth of the axioms of rational choice theory is even weaker than it was when Krantz wrote his retrospective. Many different researchers claim to have found evidence to the effect that people's choices, to a very significant extent, reflect incidental aspects of the decision situation rather than a stable, consistent preference ordering (cf. Camerer and Loewenstein 2004; Kahneman 2003). As Matthew Rabin puts it: "A lot of decisions are so sensitive to the framing or context of the choice set that it is difficult to associate these decisions as coming from framing- or context-free preferences on those choice sets" (Rabin 2002, p. 662). Similarly, Tversky writes that "it is difficult to defend the proposition that a person has a well-defined preference order (or equivalently a utility function) if different methods of elicitation give rise to different choices" (Tversky 1996, p. 189).

⁵ A reference has been omitted.

Though there are philosophers and economists who have defended orthodox economic theory, it is interesting to note that the most articulate among them in fact concede that the theory does not in general describe the observable choices of real people. Ken Binmore explicitly denies “the predictive power of economics” except when the following three conditions are simultaneously satisfied: the problem facing subjects is simple enough, incentives are large enough, and time allowed for trial-and-error adjustment is long enough (Binmore 1999, p. F17). As Binmore recognizes, this means that there is a wide range of conditions – both in the laboratory and in the field – when economic theory does not predict actual behavior. Don Ross agrees, and adds that “the proper domain of the discipline is *not* the choice behavior of individual people” (Ross 2005, p. 117). In Ross’s view, economic theory properly describes the behavior of economic *agents*, which are not identical to actual people (Ross 2005, p. 132). When the most ardent and articulate defenders of the orthodox model agree that its axioms are descriptively false of actual human beings, I trust no more elaborate argument is required.

It might be objected that the measurement-theoretic approach is adequate for the assessment of the welfare that people would have under the counterfactual condition that they are perfectly rational, fully informed, and so on, on the assumption that such counterfactual, or *ideal*, agents can be assumed to have consistent preferences. Philosophically sophisticated economists have, in fact, maintained that well-being consists in the satisfaction of those preferences that the agent would have under some specified counterfactual conditions (Harsanyi 1977, p. 646; Mongin and d’Aspremont 1998, p. 397). Yet, there is real tension between adopting the measurement-theoretic approach to measurement and adopting a preference-satisfaction account of well-being according to which what matters are counterfactual preferences. While such accounts may be more plausible as accounts of well-being, the preference ordering that one would have under some counterfactual conditions – or, the choices that one would make under those conditions – are unobservable by design. Thus, if welfare is understood in terms of the satisfaction of ideal preferences, there is no observable ordering that could serve as the basis for welfare measurement.

As long as we insist on using the measurement-theoretic approach to measurement, therefore, the best available empirical evidence strongly suggests that degrees of preference satisfaction are unmeasurable too. If well-being is a matter of ideal preferences, it is difficult to argue that there is an observable ordering that can serve as a basis for measurement. If well-being is a matter of actual preferences, it is equally difficult (but for other reasons) to argue that there is an observable ordering that can serve as a basis for measurement. Either way, the central claim underlying the argument supporting the measurability of degrees of preference satisfaction appears to be false. If mental-state accounts are inadequate because of a lack of a suitable observable ordering, preference-satisfaction accounts are inadequate for the very same reason.

However, the argument from measurability fails to undercut preference-satisfaction accounts for the same reason that it fails to undercut mental-state accounts: the proposition that measurement requires an observable ordering is false. We can be quite sure of this because the sciences get along fine without observable orderings. Contemporary psychology is dominated by psychometric methods, and this includes subfields like educational, occupational, and clinical psychology (Rust and Golombok 2009, pp. 4–5). As Robyn Dawes and Tom L. Smith point out, “the field of attitude ... is permeated by questionnaires and rating scales” (Dawes and Smith 1985, pp. 511–512). Interestingly, a

number of contemporary economists appear to have noticed the problem posed by non-standard choice behavior for welfare measurement, and have worked to make sense of the idea of welfare measurement in the absence of observable orderings satisfying the relevant axioms (Angner 2012). Thus, Jerry Green and Daniel Hojman (2007) develop a method that permits them to assess the welfare of a decision maker whether or not his or her choices satisfy the axioms of rational choice theory; Ariel Rubinstein and Yuval Salant (2008, p. 116) maintain that, given a domain of objects, economists need to distinguish what Rubinstein and Salant call “mental preference” – an unobservable “mental attitude of an individual towards the objects” – from observable choice and to develop techniques to estimate the former based on the latter using a procedure that would not rely on the usual consistency assumptions. Finally, outside of the social and behavioral sciences measurement frequently proceeds in the absence of observable orderings. Consider measurement in medicine. While an attribute like height can be measured using the measurement-theoretic approach, this is not true for most attributes of interest to medical doctors: a review of blood-pressure measurement, for example, includes no mentions of observable orderings satisfying axioms (Williams et al. 2009). It is of course in principle possible that these psychologists, economists, and medical personnel are all mistaken, and that things like attitudes, preference satisfaction, and blood pressure are impossible to measure even in principle, but I take this possibility to be too remote to be worth considering. This is not to say that there are no issues associated with measurement in psychology, economics, and medicine. Here, the point is a modest one: blood pressure measurement is possible in principle, in spite of the fact that it is not based on an observable ordering satisfying properties like transitivity.

At this point, it may be argued that I must have misrepresented the theory of measurement. Surely the presence of measurement error alone, it could be objected, must have convinced the theorists that observable orderings will not in general satisfy conditions like transitivity. According to a revisionist interpretation, measurement theory does not presuppose the existence of observable orderings satisfying the relevant axioms: when the measurement theorists were talking about “empirical relations” satisfying certain axioms, what they had in mind were unobservable – that is, what psychometricians would have called “latent” – relations. Yet, this revisionist interpretation is hard to square with the textual evidence reported in section 3. And the fact that measurement theory had no solution to the problem of measurement error, which “is always present in observation,” is a central reason Cliff (1992, p. 189) invokes to explain the failure of measurement theory. Moreover, it was intransitivities *in observable choices* – not in latent preferences – that convinced Krantz (1991) that measurement theory was unable to handle utility measurement. As one of the original authors of measurement theory, Krantz should know what it can and cannot accommodate. The revisionist interpretation fails.

To review, there are two problems with the argument from measurability. It cannot be used to build a case for the superiority of preference-satisfaction over mental-state accounts of well-being, because preferences as revealed in human choices do not in fact constitute an observable ordering of the requisite kind. Moreover, the argument cannot be used to build a case against mental-state accounts, because the proposition that measurement requires an observable ordering is false, as shown by the best scientific practice in psychology, economics,

and medicine. Again, this is not to say that there are no problems associated with mental-state accounts of well-being or with measures of happiness and satisfaction. These are the problems to which I turn next.

5 The real problems

So far, I have made the case that the argument from measurability fails: on the most charitable interpretation, the argument relies on the false notion that measurement requires the existence of an observable ordering satisfying conditions like transitivity. The failure of the argument from measurability, however, does not translate into a defense of mental-state accounts as accounts of well-being. Even if measurement does not require observable orderings, it does not follow that mental-state accounts of well-being are adequate; nor does it follow that such accounts are inadequate.⁶ Similarly, the failure of the argument from measurability does not translate into a defense of measures of happiness and satisfaction as measures of well-being. Even if measurement does not require observable orderings, it does not follow that subjectively experienced mental states can be measured, in principle or in practice, or that such measures can be used as measures of well-being; nor does it follow that such states cannot be measured or that they cannot be used as measures of well-being. Thus, the case made so far has no implications for the adequacy of mental-state accounts of well-being or the appropriateness of using measures of happiness and satisfaction as measures of well-being. In this section, I argue that there are in fact real problems associated with mental-state accounts of well-being and with measures of happiness and satisfaction, problems that may have been obscured by the ubiquity of the argument from measurability.

The paramount problem for mental-state accounts of well-being is evidence suggesting that happiness and satisfaction fail to track well-being. As James Griffin puts the problem: “The trouble with thinking of [well-being] as *one* kind of mental state is that we cannot find any one state in all that we regard as having utility – eating, reading, working, creating, helping” (Griffin 1986, p. 8). Philosophers have traditionally made the point by reference to thought experiments like Robert Nozick’s experience machine, a device that would allow superduper neuropsychologists to stimulate one’s brain so as to generate whatever experience one pleases, or evocative vignettes like Amartya Sen’s descriptions of destitute beggars, landless laborers, overworked servants, and subjugated housewives (Nozick 1974, pp. 42–43; Sen 1985, p. 15). Recently, the most consistent source of counterexamples to mental-state accounts of well-being has been the science of happiness itself. Under a wide range of conditions – from having children, being in good health and free from disability, to having ambition in life – empirical results have underscored the divergence between happiness and satisfaction on the one hand and well-being or welfare on the other. Yet, quite a number of social and behavioral scientists use words like “happiness” and “well-being” interchangeably (Angner 2011b, pp. 119–121). I have no objection to using “happiness” and “satisfaction” to denote the particular

⁶ For related reasons, the failure of the argument from measurability has no implications for the adequacy of preference-satisfaction accounts of well-being. The argument from measurability is only one argument in favor of such accounts.

subjectively experienced mental states studied by social and behavioral scientists. But it must be acknowledged that a great deal of evidence suggests that happiness and satisfaction, in this sense, fail to coincide with well-being, as the term is used in philosophical or pre-scientific literature. And this problem has nothing to do with the measurability of happiness and satisfaction.

The chief problem associated with measures of happiness and satisfaction, in my view, is that of *reification*. As *The Macmillan dictionary of psychology* defines it: “Treating an abstract idea as a real or concrete entity, a mistaken form of thinking common in children, schizophrenics, and psychologists” (Sutherland 1995, p. 392). In his classic discussion, Donald T. Campbell (1960, pp. 551–552) warns of two kinds of reification: *trait* reification, in which an abstract idea is confused with a real thing, and *score* reification, in which an actual test score is assumed to represent perfectly whatever attribute it is intended to represent. On score reification, Campbell writes: “We are all occasionally appalled at the literal interpretations and assumptions of immutable three-digit perfection with which some users regard ... test scores” (D. T. Campbell 1960, p. 552). Dawes and Smith (1985, pp. 539–540) refer to “the literal interpretation fallacy.” As they point out, the fact that a respondent checks the box next to the word “good” on a rating scale does not entail that she actually thinks of whatever she was asked to evaluate as good: “good” might simply be that category which most closely corresponds to her true attitude, she might be averse to more extreme categories, and so on. Given that reification is a well-known problem in the literature on psychometric measurement, it would be surprising if happiness researchers were immune to it.

Happiness scholars are guilty of trait reification when they confuse the latent, abstract construct – that which is represented by subjective measures – with happiness in that sense of the word that has moral significance and normative import. The two are quite different: the contemporary view of happiness as a cognitive or affective state is not what was intended when Aristotle in the *Nicomachean ethics* (1098a 16) used happiness (*eudaimonia*) to refer to “activity of the soul in conformity with excellence or virtue” (Aristotle 1962, p. 17) or when the authors of the U.S. Declaration of Independence declared “the pursuit of happiness” to be an inalienable right (Haybron 2000). The process of construct validation at best establishes the existence of *some* construct represented by the relevant measure; even when successful, no process of construct validation can by itself establish that the underlying construct has anything to do with “happiness” as the term is used in philosophical or pre-scientific literature.

And yet, happiness in the sense of modern psychology and economics, on the one hand, and happiness as it is used by in the non-scientific literature, on the other, are often confused. Angus Campbell explains his interest in happiness as follows:

The attraction of the concept of happiness is certainly great, coming as it does from the early Greek identification of happiness with the good life and having as it does almost universal currency as a recognized, if not uniquely important, component of the quality of life experience (A. Campbell 1976, p. 119).

Tal Ben-Shahar (2007, p. 31) opens a chapter called “Happiness Explained” with an epigraph from Aristotle: “Happiness is the meaning and the purpose of life, the whole aim and end of human existence.” Lyubomirsky and Lepper write that “tremendous

interest” in happiness can be found from “Aristotle and the writers of the American Declaration of Independence to present-day philosophers, politicians, novelists, and authors of popular psychology” (Lyubomirsky and Lepper 1999, p. 137). By equivocating in this manner, these passages falsely suggest that the term “happiness” as it is used by the author is the same as what non-scientific literature would consider “true” happiness, and in this sense conflate the abstract, latent construct with the real thing. Notice that the equivocation appears in peer-reviewed literature as well as in popular writings. Trait reification is pernicious, because it obscures the meaning of empirical results and exaggerates their significance.

Happiness researchers are guilty of score reification and the literal-interpretation fallacy when they report more significant figures than are warranted by the data and when they infer that people’s attitudes correspond perfectly to their answers on a rating scale. Paul Frijters and his co-authors make themselves open to the charge of score reification when they draw conclusions about the “birth of a child being worth the equivalent of a windfall gain of \$238,000” (Frijters et al. 2011, p. 207), since the statement gives the impression of the “immutable three-digit perfection” that Campbell deplored. Richard Layard may be guilty of the literal-interpretation fallacy when he writes that “some 45% of the richest quarter of Americans are very happy, compared with only 33% of the poorest quarter” (Layard 2005, p. 30), since the statement appears to be a literal interpretation of responses to the question of Gurin et al. (1960). So too for Ed Diener and Carol Diener (1996), who review similar evidence and announce: “Most people are happy.” Ironically, Gurin et al. themselves warned against this fallacy. After noting that in their study “11 per cent of those interviewed said ‘not too happy,’” they added: “One should avoid interpreting this response as an absolute statement that ‘11 per cent of the American people are unhappy’” (Gurin et al. 1960, p. 15). Notice, again, that the equivocation appears in peer-reviewed literature as well as in popular writings. Score reification and the literal-interpretation fallacy are pernicious, because they obscure the imprecision inherent in empirical findings and therefore might lead to bad science and policy.

In this section, I have argued that the failure of the argument from measurability does not translate into a defense of mental-state accounts of well-being or of subjective measures of well-being. I have also argued that the ubiquity of the argument from measurability may have obscured real problems with mental-state accounts of well-being – above all, that happiness and satisfaction fail to track well-being – and with measures of happiness and satisfaction – above all, the tendency toward reification. Happiness scholars have made themselves open toward the charge of both trait and score reification. In fact, the tendency toward trait reification may have been a driving force behind the equivocation between happiness in the sense of a subjectively experienced mental state and happiness or well-being in the sense that has moral significance and normative import. Notice, though, that reification is a problem associated with particular uses of happiness and satisfaction measures, and not with the measures per se. It is perfectly possible to use subjective measures without engaging in reification.

6 Conclusion

In this paper, I have articulated and assessed the “argument from measurability,” as it appears in the writings of philosophers and economists critical of mental-state accounts of well-being. In order to assess the argument, I reconstructed it – as hinted at by authors like Chappell and Crisp, Fehige and Wessels, and Beckerman – as charitably as possible. I offered multiple reasons to think that my reconstruction is true to the letter and spirit of the relevant writings. A critic could still insist that I have failed to properly reconstruct the argument, but to substantiate the charge the critic would need to offer a superior reconstruction. On my interpretation, which I will favor in the absence of a superior one, the argument depends critically on the proposition that measurement requires the existence of an observable ordering satisfying conditions like transitivity. If this proposition were true, it would be an embarrassment not just for mental-state accounts of well-being but for preference-satisfaction accounts as well, since neither actually observed (revealed) preferences or counterfactual (ideal) preferences constitutes an observable ordering of the requisite kind. In reality, however, actual scientific practice strongly suggests that measurement does not require the existence of an observable ordering satisfying conditions like transitivity. On the most charitable interpretation, then, the argument from measurability fails because it proceeds from a false premiss.

Some caveats are in order. First, I do not pretend that the argument from measurability is the only argument against mental-state accounts of well-being, as Chappell and Crisp seem to suggest. It deserves our attention because it is so common, but other arguments (like that hinted at in section 5) may ultimately be more successful. Second, even if I am right that the argument from measurability is unsound, I do not claim to have offered a conclusive positive argument to the effect that subjectively experienced mental states like happiness and satisfaction can in principle be measured; nor have I argued that degrees of preference satisfaction cannot be measured. Third, even if it has been shown that mental states like happiness and satisfaction can be measured in principle, it is not at all obvious that they can be measured in practice. Problems of reliability and validity of subjective measures are well known to psychologists. As Diener et al. write: “SWB values may change depending on the type of scales used, the order of items, the time-frame of the questions, current mood at the time of measurement, and other situational factors” (Diener et al. 1999, p. 278).⁷ In the light of phenomena like these, Norbert Schwarz and Fritz Strack write: “Reports of subjective well-being (SWB) do not reflect a stable inner state of well-being. Rather, they are judgments that individuals form on the spot, based on information that is chronically or temporarily accessible at that point in time, resulting in pronounced context effects” (Schwarz and Strack 1999, p. 61). The degree of validity and reliability are of course empirical questions, which cannot be settled by philosophical reflection alone. Fourth, as discussed in section 5, the rejection of the argument from measurability does not translate into a defense of mental-state accounts as accounts of well-being or of happiness and satisfaction measures as measures of well-being.

The ubiquity of the measurement-theoretic argument may have obscured very real problems that have nothing to do with the existence of such orderings. Chief among

⁷ References have been omitted.

the problems associated with happiness measurement, I have suggested, is the problem of reification. Given that reification is a well-known problem in the literature on psychometric measurement, it is unsurprising that happiness researchers should be prone to it. Happiness scholars make themselves open to the charge of both trait and score reification, as they appear to confuse the construct they are studying with happiness in that sense of the word which has moral significance and normative import, report more significant figures than are warranted by the data, and infer that people's attitudes correspond perfectly to their answers on a rating scale. Reification is pernicious, because it obscures the meaning of empirical results, exaggerates their significance, obscures the imprecision inherent in them, and consequently leads to bad science and policy.

The central problem associated with happiness measurement, then, is not that happiness is *too hard* to measure, as it would be if the argument from measurability were sound. Rather, there is a sense in which the problem is the opposite. Distributing questionnaires and plugging the resulting data into widely accessible software packages in order to generate statistical estimates is simple indeed, and the results (which often are generated with a precision of eight decimal places or more) are easily misinterpreted. If it is in fact true that happiness can be validly measured using as little as one direct question, doing so is as easy as measuring anything in the social and behavioral sciences, and considerably easier than measuring, e.g., depression using the 21-item Beck Depression Inventory (Beck et al. 1961) or health status using the 36-item Short Form Health Survey (SF-36) (Ware and Sherbourne 1992). And while reification is a potential problem for all measurement of latent constructs, there is reason to think that the problem is particularly grave in the case of happiness and satisfaction measures. The reason is the ease with which psychologists and economists draw conclusions about happiness or well-being in that sense of the word that has moral significance and normative import and offer public policy prescriptions on that basis. In the sense that the ease with which data can be gathered and analyzed encourages the tendency toward reification, then, the problem is rather that happiness in the sense of a subjectively experienced mental state is *too easy* to measure.

Acknowledgment I am grateful to Zvi Biener, Robyn Dawes, Greg Frost-Arnold, Daniel Hausman, Brian Hepburn, Harold Kincaid, Peter Machamer, Gualtiero Piccinini, Nicholas Rescher, Don Ross, Sam Wren-Lewis, and anonymous referees for constructive comments on earlier drafts. Errors remain my own.

References

- Alexandrova, A. (2008). First-person reports and the measurement of happiness. *Philosophical Psychology*, 21(5), 571–583. doi:10.1080/09515080802412552.
- Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: Americans' perceptions of life quality*. New York: Plenum Press.
- Angner, E. (2009a). The politics of happiness: Subjective vs. economic measures as measures of social well-being. In L. Bortolotti (Ed.), *Philosophy and happiness* (pp. 149–166). New York: Palgrave Macmillan.
- Angner, E. (2009b). Subjective measures of well-being: Philosophical perspectives. In H. Kincaid & D. Ross (Eds.), *The Oxford handbook of philosophy of economics* (pp. 560–579). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780195189254.003.0021.

- Angner, E. (2010). Subjective well-being. *Journal of Socio-Economics*, 39(3), 361–368. doi:10.1016/j.socsec.2009.12.001.
- Angner, E. (2011a). The evolution of eupathics: The historical roots of subjective measures of wellbeing. *International Journal of Wellbeing*, 1(1), 4–41. doi:10.5502/ijw.v1i1.14.
- Angner, E. (2011b). Are subjective measures of well-being “direct”? *Australasian Journal of Philosophy*, 89(1), 115–130. doi:10.1080/00048400903401665.
- Angner, E. (2012). Current trends in welfare measurement. In J. B. Davis & D. W. Hands (Eds.), *The Elgar companion to recent economic methodology* (pp. 121–154). Northampton: Edward Elgar.
- Angner, E., & Loewenstein, G. (2012). Behavioral economics. In U. Mäki (Ed.), *Handbook of the philosophy of science: Philosophy of economics* (pp. 641–690). Amsterdam: Elsevier.
- Aristotle (1962). *Nicomachean ethics*. (M. Oswald, Trans.). Indianapolis, IN: Bobbs-Merrill Co.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4(6), 561–571. doi:10.1001/archpsyc.1961.01710120031004.
- Beckerman, W. (1975). *Two cheers for the affluent society: A spirited defense of economic growth*. New York: Saint Martin's Press.
- Ben-Shahar, T. (2007). *Happier: Learn the secrets to daily joy and lasting fulfillment*. New York: McGraw-Hill.
- Binmore, K. (1999). Why experiment in economics? *The Economic Journal*, 109(453), F16–F24. doi:10.1111/1468-0297.00399.
- Brooks, A. C. (2008). *Gross National Happiness: Why happiness matters for America—and how we can get more of it*. New York: Basic Books.
- Camerer, C. F., & Loewenstein, G. (2004). Behavioral economics: Past, present, future. In C. F. Camerer, G. Loewenstein, & M. Rabin (Eds.), *Advances in behavioral economics* (pp. 3–51). New York and Princeton: Russell Sage Foundation and Princeton University Press.
- Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist*, 15(8), 546–553. doi:10.1037/h0048255.
- Campbell, A. (1976). Subjective measures of well-being. *American Psychologist*, 31(2), 117–124. doi:10.1037/0003-066X.31.2.117.
- Cantril, H. (1965). *The pattern of human concerns*. New Brunswick: Rutgers University Press.
- Chappell, T., & Crisp, R. (1998). Utilitarianism. In E. Craig (Ed.), *Routledge encyclopedia of philosophy* (Vol. 9, pp. 551–557). London: Routledge.
- Cliff, N. (1992). Abstract measurement theory and the revolution that never happened. *Psychological Science*, 3(3), 186–190. doi:10.1111/j.1467-9280.1992.tb00024.x.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. doi:10.1037/h0040957.
- Dawes, R., & Smith, T. L. (1985). Attitude and opinion measurement. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 1, pp. 509–566). New York: Random House.
- Diener, E. (2000). Subjective well-being. The science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34–43. doi:10.1037/0003-066X.55.1.34.
- Diener, E. (2006). Guidelines for national indicators of subjective well-being and ill-being. *Applied Research in Quality of Life*, 1(2), 151–157. doi:10.1007/s11482-006-9007-x.
- Diener, E., & Diener, C. (1996). Most people are happy. *Psychological Science*, 7(3), 181–185. doi:10.1111/j.1467-9280.1996.tb00354.x.
- Diener, E., & Seligman, M. E. P. (2004). Beyond money: Toward an economy of well-being. *Psychological Science in the Public Interest*, 5(1), 1–31. doi:10.1111/j.0963-7214.2004.00501001.x.
- Diener, E., & Suh, E. M. (1997). Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research*, 40(1–2), 189–216. doi:10.1023/A:1006859511756.
- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125(2), 276–302. doi:10.1037/0033-2909.125.2.276.
- Diener, E., Lucas, R., Schimmack, U., & Helliwell, J. (2009). *Well-being for public policy*. New York: Oxford University Press.
- Fehige, C., & Wessels, U. (1998). Preferences: An introduction. In C. Fehige & U. Wessels (Eds.), *Preferences* (pp. xx–xlili). Berlin: W. de Gruyter.
- Frey, B. S., & Stutzer, A. (2000). Maximising happiness? *German Economic Review*, 1(2), 145–167. doi:10.1111/1468-0475.00009.
- Frijters, P., Johnston, D. W., & Shields, M. A. (2011). Life satisfaction dynamics with quarterly life event data. *The Scandinavian Journal of Economics*, 113(1), 190–211. doi:10.1111/j.1467-9442.2010.01638.x.
- Gilbert, D. T. (2006). *Stumbling on happiness*. New York: A.A. Knopf.

- Green, J., & Hojman, D. (2007). *Choice, rationality and welfare measurement*. Harvard Institute of Economic Research Discussion Paper No. 2144; KSG Working Paper No. RWP07-054. Version of November 2007. Retrieved from <http://ssrn.com/abstract=1030342>
- Griffin, J. (1986). *Well-being: Its meaning, measurement, and moral importance*. Oxford: Clarendon.
- Gurin, G., Veroff, J., & Feld, S. (1960). *Americans view their mental health: A nationwide interview survey*. New York: Basic Books.
- Harsanyi, J. C. (1977). Morality and the theory of rational behavior. *Social Research*, 44(4), 623–656.
- Hart, H. (1940). *Chart for happiness*. New York: Macmillan.
- Haybron, D. M. (2000). Two philosophical problems in the study of happiness. *Journal of Happiness Studies*, 1(2), 207–225. doi:10.1023/A:1010075527517.
- Haybron, D. M. (2008). *The pursuit of unhappiness: The elusive psychology of well-being*. Oxford: Oxford University Press.
- John, O. P., & Benet-Martínez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). New York: Cambridge University Press.
- Johnson, J. A. (2001). Personality psychology: Methods. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 16, pp. 11313–11317). Oxford: Pergamon. doi:10.1016/B0-08-043076-7/01671-5.
- Jones, L. V., & Thissen, D. (2006). A history and overview of psychometrics. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, pp. 1–27). Amsterdam: North-Holland.
- Judd, C. M., & McClelland, G. H. (1998). Measurement. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 180–232). New York: Oxford University Press.
- Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York: Russell Sage.
- Kahneman, D. (2003). A psychological perspective on economics. *American Economic Review*, 93(2), 162–168. doi:10.1257/000282803321946985.
- Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1), 3–24. doi:10.1257/089533006776526030.
- Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004a). A survey method for characterizing daily life experience: The Day Reconstruction Method. *Science*, 306(5702), 1776–1780. doi:10.1126/science.1103572.
- Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. (2004b). Toward national well-being accounts. *American Economic Review*, 94(2), 429–434. doi:10.1257/0002828041301713.
- Krantz, D. H. (1991). From indices to mappings: The representational approach to measurement. In D. R. Brown & J. E. K. Smith (Eds.), *Frontiers of mathematical psychology: Essays in honor of Clyde Coombs* (pp. 1–52). New York: Springer.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Additive and polynomial representations* (Vol. 1). New York: Academic.
- Krueger, A. B. (Ed.). (2009). *Measuring the subjective well-being of nations: National accounts of time use and well-being*. Chicago: The University of Chicago Press.
- Layard, P. R. G. (2005). *Happiness: Lessons from a new science*. New York: Penguin Press.
- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, 46(2), 137–155. doi:10.1023/A:1006824100041.
- Mandler, M. (1999). *Dilemmas in economic theory: Persisting foundational problems of microeconomics*. New York: Oxford University Press.
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. New York: Oxford University Press.
- Mongin, P., & D'Aspremont, C. (1998). Utility theory and ethics. In S. Barberà, P. J. Hammond, & C. Seidl (Eds.), *Handbook of utility theory* (Vol. 1, pp. 371–481). Dordrecht: Kluwer Academic Publishers.
- Nozick, R. (1974). *Anarchy, state, and utopia*. New York: Basic Books.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Nussbaum, M. C. (2008). Who is the happy warrior? Philosophy poses questions to psychology. *The Journal of Legal Studies*, 37(S2), S81–S113. doi:10.1086/587438.
- Piccinini, G. (2009). First person data, publicity and self-measurement. *Philosophers' Imprint*, 9(9), 1–16.
- Pigou, A. C. (1952). *The economics of welfare* (4th ed.). London: Macmillan.
- Rabin, M. (2002). A perspective on psychology and economics. *European Economic Review*, 46(4–5), 657–685. doi:10.1016/S0014-2921(01)00207-0.
- Rabinowicz, W., & Österberg, J. (1996). Value based on preferences. *Economics and Philosophy*, 12(1), 1–27. doi:10.1017/S0266267100003692.

- Ross, D. (2005). *Economic theory and cognitive science: Microexplanation*. Cambridge: MIT Press.
- Rubinstein, A., & Salant, Y. (2008). Some thoughts on the principle of revealed preference. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 115–124). New York: Oxford University Press.
- Rust, J., & Golombok, S. (2009). *Modern psychometrics: The science of psychological assessment* (3rd ed.). New York: Routledge.
- Scanlon, T. (2000). *What we owe to each other*. Cambridge: Belknap.
- Schwartz, N., & Strack, F. (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 61–84). New York: Russell Sage.
- Scott, D., & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic*, 23(2), 113–128. doi:[10.2307/2964389](https://doi.org/10.2307/2964389).
- Sen, A. (1985). *Commodities and capabilities*. Amsterdam: North-Holland.
- Sen, A. (2002). *Rationality and freedom*. Cambridge: Belknap.
- Sumner, L. W. (1996). *Welfare, happiness, and ethics*. Oxford: Clarendon.
- Sutherland, N. S. (1995). *The Macmillan dictionary of psychology* (2nd ed.). Basingstoke: Macmillan.
- Tiberius, V., & Plakias, A. (2010). Well-being. In J. M. Doris & The Moral Psychology Research Group (Eds.), *The moral psychology handbook* (pp. 402–431). Oxford: Oxford University Press.
- Trout, J. D. (1998). *Measuring the intentional world: Realism, naturalism, and quantitative methods in the behavioral sciences*. New York: Oxford University Press.
- Tversky, A. (1996). Rational theory and constructive choice. In K. J. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The rational foundations of economic behaviour: Proceedings of the IEA conference held in Turin, Italy* (pp. 185–197). Basingstoke: Macmillan Press.
- Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 30(6), 473–483.
- Watson, G. (1930). Happiness among adult students of education. *Journal of Educational Psychology*, 21(2), 79–109. doi:[10.1037/h0070539](https://doi.org/10.1037/h0070539).
- Williams, J. S., Brown, S. M., & Conlin, P. R. (2009). Blood-pressure measurement. *The New England Journal of Medicine*, 360(5), e6–e9. doi:[10.1056/NEJMvcm0800157](https://doi.org/10.1056/NEJMvcm0800157).
- Wilson, W. (1967). Correlates of avowed happiness. *Psychological Bulletin*, 67(4), 294–306. doi:[10.1037/h0024431](https://doi.org/10.1037/h0024431).