SUBJECTIVE MEASURES OF WELL-BEING

A PHILOSOPHICAL EXAMINATION

by

Erik Angner

B.A., Uppsala University, 1995

M.A., Uppsala University, 1997

Ph.D., University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

FAS in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Erik Angner

It was defended on

June 8, 2005

and approved by

Henry Krips, Professor, Communication (University of Pittsburgh)

George Loewenstein, Professor, Social and Decision Sciences (Carnegie Mellon University)

Nicholas Rescher, University Professor, Philosophy (University of Pittsburgh)

Peter Machamer, Professor, History and Philosophy of Science (University of Pittsburgh)
Dissertation Director

SUBJECTIVE MEASURES OF WELL-BEING

A PHILOSOPHICAL EXAMINATION

Erik Angner, PhD

University of Pittsburgh, 2005

Over the last couple of decades, as part of the rise of positive psychology, psychologists have given increasing amounts of attention to so-called *subjective measures of well-being*. These measures, which are supposed to represent the well-being of individuals and groups, are often presented as alternatives to more traditional economic ones for purposes of the articulation, implementation and evaluation of public policy. Unlike economic measures, which are typically based on data about income, market transactions and the like, subjective measures are based on answers to 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?" The aim of this dissertation is to explore issues in the philosophical foundations of subjective measures of well-being, with special emphasis on the manner in which the philosophical foundations of subjective measures differ from those of traditional economic measures. Moreover, the goal is to examine some arguments for and against these measures, and, in particular, arguments that purport to demonstrate the superiority of economic measures for purposes of public policy. My main thesis is that the claim that subjective measures of well-being cannot be shown to be inferior to economic measures quite as easily as some have suggested, but that they nevertheless are associated with serious problems, and that questions about the relative advantage of subjective and economic measures for purposes of public policy will depend on some fundamentally philosophical judgments, e.g. about the nature of well-being and the legitimate goals for public policy.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

This dissertation is the result of the wonderfully stimulating interdisciplinary environment in Pittsburgh, where my work was shaped by innumerable conversations around the Department of History and Philosophy of Science, the Department of Philosophy, and the Department of Economics at the University of Pittsburgh, as well as the Department of Social and Decision Sciences and the Department of Philosophy at Carnegie Mellon University. It is my hope that the result fruitfully incorporates and synthesizes insights from the various fields represented by these departments.

Among the individuals who deserve special mention are Prof. Peter Machamer, who very kindly agreed to supervise this project and who remained supportive in spite of its unpredictable development, as well as Profs. Henry Krips, George Loewenstein, and Nicholas Rescher, who very competently served as members of the committee. I am deeply grateful to all of the above for their penetrating insights and constructive criticism. As always, neither of them should be held responsible for the errors that remain.

Finally, I wish to express my gratitude for generous extra-departmental financial support, without which the ambition to simultaneously pursue two doctoral degrees would have been positively foolish. Chief among them is Jan Wallander and Tom Hedelius' Foundation, whose munificent Hedelius Fellowship supported what in retrospect appear as the two most creative years of my graduate studies. I also gratefully acknowledge the support of The Sweden America Foundation, *Borgrättsfonderna* at the Office of the Marshal of the Realm (Sweden), *Byzantinska resefonden* at Uppsala University, *Carl Erik Levins Stiftelse* at *Skandinaviska Enskilda Banken*, the Institute

for Humane Studies, the *Fondazione Famiglia Rausing* at the Swedish Institute in Rome, as well as the FAS–PBC Fellowship Graduate Fund and the Dean's Tuition Scholarship Fund at the University of Pittsburgh.

The dissertation is dedicated to Elizabeth.

1.0 INTRODUCTION

Abstract. In this chapter, I introduce the topic of subjective measures of well-being, provide some historical context, and clarify the nature of my project.

1.1 INTRODUCTION

The concept of well-being plays a prominent role in a number of disciplines. It appears not only in various subfields of philosophy – especially in ethics and political philosophy – but also in economics, psychology, psychiatry, public health, gerontology, and elsewhere. Although the exact function of the concept varies across disciplines and across authors, it is typically expected to play several extraordinarily important roles. Thomas Scanlon (1998) articulates these roles as follows:

> It is commonly supposed that there is a simple notion of individual well-being that plays the following three roles. First, it serves as an important basis for the decisions of a single rational individual, at least for those decisions in which he or she alone is concerned (that is to say, in which moral obligations and concerns for others can be left aside). Second, it is what a concerned benefactor, such as a friend or parent, has reason to promote. Third, it is the basis on which an individual's interests are taken into account in moral argument (Scanlon 1998, 93).[1]

Whatever the exact function of the concept of well-being in various disciplines, it certainly is often assumed to play a role in determining both what I should pursue in my own life, and what I should

[1] Scanlon proceeds to criticize the view outlined in this passage. This does not change the fact that it is a nice characterization of the role that the concept of well-being is often supposed to play. We will come back to the topic in chapter 3.0.

promote in the lives of others. Incidentally, the concept of well-being is often applied to groups and nations as well as to individuals, and supposed to play as a basis for the deliberations by governments regarding public policy.

Perhaps unsurprisingly, the literature refers to this "simple notion" well-being in a variety of ways. For example, in his book *Welfare, Happiness and Ethics*, L. W. Sumner (1996) writes that "a person's welfare is more or less the same as her well-being or interest or (in one of its many meanings) her good" (Sumner 1996, 1). Similarly, in the words of Andrew Moore and Roger Crisp: "At a minimum, a life of well-being is a life going well. The numerous near-equivalents to well-being include a person's good, benefit, advantage, interest, prudential value, welfare, happiness, flourishing, *eudaimonia*, and utility" (Moore and Crisp 1996, 599). Other terms that could have been added to this list include "quality of life" and "thriving" (see e.g. Nussbaum and Sen 1993, 1). Simon Keller writes:

> I will treat "what advances your welfare," "what makes you better off," "what makes your life go well" and "what's in your best interests" as synonymous, while being aware that these phrases may have slightly varying connotations in ordinary language. In any case, my subject is the notion of welfare or well-being discussed in, for example, [Parfit (1984), Griffin (1986), and Sumner (1996)] (Keller 2004, 39).

Incidentally, passages like these support Scanlon's contention that the different terms are typically used to denote one "simple notion" of well-being rather than a multiplicity of related notions.[2]

Equally unsurprisingly, there have been many attempts to develop adequate measures of well-being.[3] Very often, these attempts are motivated by a desire to help governments and other decision makers design policy so as to promote people's well-being. Although it can be argued that

[2] This is not to say that there are not other, related notions that need to be distinguished from "well-being" as the term is understood here. See chapter 3 for a fuller discussion of these concepts.

[3] For a longer discussion of this topic, see section 1.2 below.

such attempts have a much longer history, a prominent effort in this direction is evident in the work of A. C. Pigou, who is commonly considered the father of welfare economics. In *The Economics of Welfare* (1960 [1920]), Pigou was explicit about his desire to develop a measure of welfare that could be used in practice. In his own words, the goal was "to make more easy practical measures to promote welfare – practical measures which statesmen may build upon the work of the economist" (Pigou 1960, 10). Later on, the social indicator movement developed measures based on a broader range of statistics, including life expectancy, access to health care, housing conditions, and so on, and argued that these composite measures were superior to the economic ones for policy purposes (cf. Campbell 1976).

This dissertation deals with so-called subjective measures of well-being.[4] These measures are intended to reflect mental states such as happiness, satisfaction, engagement, and so on, and they are typically presented as alternatives to traditional economic measures and social indicators for purposes of public policy. These measures, and the empirical generalizations that have been defended by reference to them – sometimes referred to as the "science of happiness"[5] – have attracted a great deal of attention in recent years. At the time of writing, both *Psychology Today* and *Time Magazine* have just run cover stories on happiness research, and the topic has been discussed recently by among others *Wired*, *The Financial Times*, and NBC. Moreover, it is not just the popular press that pays attention. Recent Nobel laureate Daniel Kahneman (who has been one of the main proponents of the new measures) and co-authors had an article on the topic in a December 2004 issue of *Science*.

The subjective measures have also met fierce resistance, especially from economists who are comfortable using traditional measures. However, arguments both for and against subjective

[4] For a longer discussion, see section 1.3 below.

[5] See the cover story of the January 17, 2005, issue of *Time Magazine*.

measures are often difficult to assess. First, as they appear in the literature, these arguments are typically incomplete. Often, important premisses have been suppressed, and it is not evident what those premisses are. Uncertainty on this score makes it difficult to assess both the truth of the premisses and the validity of the argument. Second, the philosophical foundations of subjective measures remain unclear. For example, as we will see below (in chapters 2.0 and 3.0), it is not always clear even what the psychologists mean by "well-being." Presumably, this second fact helps explain the first; if the philosophical assumptions underlying a certain approach are obscure, it is hard to develop clear and explicit arguments for and against them. Anyway, while this lack of clarity about the philosophical foundations may be unsurprising in light of the relative youth of the literature, it also makes it hard to assess arguments.

The aim of this dissertation is to explore the philosophical foundations of subjective measures of well-being, and to examine some arguments that have been offered against them. My hope is that exploring the foundations of subjective measures will help us articulate, and assess the soundness of, different arguments for and against these measures. Since subjective measures are often presented as alternatives to traditional economic measures, I will focus on the ways in which the foundations of subjective measures differ from those of the more widely used economic measures of welfare, and on arguments that purport to demonstrate the superiority of economic measures for purposes of public policy.

Because there are so many open questions regarding the philosophical assumptions of subjective measures, a large part of the dissertation will be dedicated to exploring the assumptions that underlie the psychologists' approach to the measurement of well-being, and how they relate to the economists'. In practice, it is often impossible to identify all the relevant assumptions that go into an argument simply by reading the articles in which the argument is presented. In addition to examining the articles of the relevant psychologists and economists, therefore, I will also examine

the history of their disciplines. The hypothesis is that we can develop a fuller and more accurate picture about the nature of an intellectual enterprise by tracing its historical roots.

My main thesis is that subjective measures of well-being cannot be shown to be inferior to economic measures quite so easily as some have suggested, but that they nevertheless are associated with serious problems. Moreover, I claim, questions about the relative advantage of subjective and economic measures for purposes of public policy will depend on some fundamentally philosophical judgments, e.g. about the nature of well-being and the legitimate goals for public policy. Thus, this dissertation can be read as a qualified defense of subjective measures. It is a *defense* because I suggest that many of the arguments that have been presented against subjective measures (and for economic measures) fail to achieve their goal. It is a *qualified* defense because I argue that subjective measures are associated with so many serious problems that it remains unclear if they represent an improvement over traditional economic measures.

1.2 THE SETTING

In this section, I offer a working definition of measurement, and a brief history of attempts to measure well-being. (An extended discussion of the meaning of "well-being" appears in section 3.2). The idea is to set the stage for the discussion that follows by putting subjective measures in historical context.

1.2.1 The idea of measurement

As a working definition of "measurement," I will adopt the useful and representative definition articulated by David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky in their

monumental *Foundations of Measurement* (1971). Krantz et al. write: "When measuring some attribute of a class of objects or events, we associate numbers (or other familiar mathematical entities, such as vectors) with the objects in such a way that the properties of the attribute are faithfully represented as numerical properties" (Krantz et al. 1971, 1).[6] Thus, loosely speaking, *measurement* is the process of assigning numbers of objects so as to represent some property, and a *measure* is simply a function from a set of objects e.g. to a set of numbers. Note that a measure is not an operational definition in the sense of Bridgman (1927). The properties of interest are typically assumed to be defined independently of the measure used to represent them (cf. Ghiselli et al. 1981, 15). In the context of the measurement of well-being, the property in questions is, of course, well-being, and the objects are individuals. Sometimes various measures are used to represent the well-being of groups, though I will largely ignore the additional complications that arise in such contexts.

1.2.2 Economic measures

Efforts by economists to measure well-being (or welfare) goes back at least to Pigou's book *The Economics of Welfare* (Pigou 1960 [1920]), which marks the beginning of the field of welfare economics (Hicks 1975, 307). In the words of John C. Chipman and James C. Moore, "Pigou's object was quite explicitly to obtain an index of welfare" (Chipman and Moore 1976, 391). According to Pigou himself, the goal is "to make more easy practical measures to promote welfare – practical measures which statesmen may build upon the work of the economist" (Pigou 1960, 10). Pigou writes:

[6] Cf. Allen and Yen (1979, 2), Roberts (1979, 49-50), and Nunnally and Bernstein (1994, 1).

> The one obvious instrument of measurement available in social life is money. Hence, the range of our inquiry becomes restricted to that part of social welfare that can be brought directly or indirectly into relation with the measuring-rod of money. This part of welfare may be called economic welfare (Pigou 1960, 11).

The specific measure that Pigou proposes is the *national dividend*, or national income, by which he means "that part of the objective income of the community, including, of course, income derived from abroad, which can be measured in money" (Pigou 1960, 31). The use of the national dividend as a measure of welfare is justified, in Pigou's view, by the belief that the size of the dividend is highly correlated with the degree of economic welfare of the nation. He writes: "The economic welfare of the country is intimately associated with the size of the national dividend, and changes in economic welfare with changes in the size of the dividend" (Pigou 1960, 50).

Concepts related to the national dividend, including that of real income, remain some of the most commonly used measures of welfare. Donald Rutherford writes that the social welfare of a country is "often measured by the total volume of goods and services becoming available to it over a given period, i.e. real income" (Rutherford 2002, 521). Martha Nussbaum and Amartya Sen point out that measures like Gross National Product (GNP) per capita, in spite of their well-known shortcomings, "continue to be widely used when public policy is made" (Nussbaum and Sen 1993, 2). The widespread concern with economic growth also testifies to the importance of real income as a measure of well-being. Since "growth" is often used to refer to the first derivative of the national product, and "growth rate" to refer to the second derivative, high growth (or a high growth rate) can be seen as an indication of more well-being in the future.

Measuring national income in practice is, of course, a non-trivial problem, especially at the national level. The figure depends on a number of arbitrary decisions, and the data is often incomplete or unreliable. Beckerman – who, incidentally, is the author of *In Defense of Economic Growth* (1974) – writes:

> It cannot be too strongly emphasized that any figure of GNP or of national income involves an enormous amount of estimation on the basis of what are often very shaky assumptions and inadequate data. As a result, national income estimates are frequently subject to very large revisions.... It is also quite common for there to be more than one quasi-official set of estimates of GNP in the same country (Beckerman 1987, 591).

Thus, there are many reasons to distrust the reliability of estimates of GDP and GNP.

An alternative way to evaluate the welfare consequences of policy interventions is in terms of *consumer surplus* (CS) and *producer surplus* (PS). The notion of a consumer surplus goes back to Jules Dupuit (1969 [1844]). Dupuit was concerned with spelling out the conditions under which public works – such as the building of a bridge – can "be declared of public utility" (Dupuit 1969, 255). He writes: "Political economy has to take as the measure of the utility of an object the maximum sacrifice which each consumer would be willing to make in order to acquire the object" (Dupuit 1969, 262). According to Dupuit's method, we can calculate the public utility of a project by adding up such utilities. The bridge should be built, he argued, if the public utility (in francs) exceeds the costs of construction. Dupuit's idea was further developed, and popularized, by Alfred Marshall (1948 [1890]). Marshall defined consumer surplus of a good as "[the] excess of the price which [the consumer] would be willing to pay rather than go without the thing, over that which he actually does pay" (Marshall 1948, 124).

As Marshall suggests, the consumer surplus is given by the difference between the maximum amount of money an agent would pay for a good and the amount she actually paid for it. Assume that I would pay at most 5 dollars for my first widget, 3 dollars for the second, and 1 for the third. If the price of widgets is 2 dollars, then, I would buy two widgets. The consumer surplus in this case is the sum of the difference between my willingness-to-pay for each unit and the actual price, i.e. (5-2)+(3-2) = 4 dollars. Equivalently, we can compute the consumer surplus by taking the sum of the willingness-to-pay for each unit of the good bought and subtract the total expenditure, i.e. (5+3)-

(2*2) = 4 dollars. In geometrical terms, we plot the willingness-to-pay as a function of the quantity. The curve will then coincide with the demand curve for the good. Now, the consumer surplus is given by the area below the demand curve but above the price line (Just et al. 1982, 72).

The notions of consumer and producer surplus – and *total surplus*, the sum of the two – are widely used to evaluate the consequences of public policy. According to Daniel T. Slesnick: "Consumer surplus is the overwhelming choice as a welfare indicator" (Slesnick 1998, 2110). Just et al. appear to agree: "'Consumer surplus' is the vehicle most often used in empirical work to measure consumer welfare" (Just et al. 1982, 69-70). Moreover, surplus is the tool preferred by undergraduate textbooks when evaluating the welfare consequences of interventions like price ceilings and trade restrictions (cf. Mankiw 2001, 'Part III: Markets and Welfare').

How, in practice, do economists go about calculating measures of consumer and surplus? As we have seen, this calculation requires an estimation of the demand and supply curves. Estimating the shape of demand and supply curves is a task for econometrics.[7] Just et al. (1982, 165-173) discuss this issue in the case of market goods. They write: "The usual approach in econometrics is to assume a particular functional form that is, hopefully, sufficiently general to describe adequately the mechanism generating a particular set of data" (Just et al. 1982, 166-167). Having picked a particular functional form, the econometrician may calculate those parameters "that minimize the sum of squared deviations of the observed quantities from the estimated linear relationship" (Just et al. 1982, 167). Using the result as an estimation of the true parameters, it is easy enough to produce an estimate of the curve as a whole, and using our estimate of the curve we can compute the consumer, producer or total surplus associated with a change (Just et al. 1982, 168). However, estimated demand and supply curves depend crucially on the availability of data, assumptions about the

[7] Rutherford defines econometrics as follows: "The measurement of economic relationships using statistical techniques, and the testing of economic theories" (Rutherford 2002, 158).

functional form, among other things (cf. Just et al. 1982, 171-173). As a result, estimates are often inexact.

Yet another set of measures evolve around the concepts of *compensating variation* (CV) and *equivalent variation* (EV). These notions were developed in a series of publications by John R. Hicks (1941; 1942; 1943), and are neatly defined by Just et al. (1982). Assuming we are interested in the welfare change for a consumer going from one state *A* to another state *B*, the compensating variation is the amount of money that one would need to take away from a consumer in state *B* to restore her utility level of state *A* (Just et el. 1982, 85). Similarly, the equivalent variation is the amount of money that one would have to give to the consumer in state *A* to leave her with the same utility level that she would get if she moved to state *B* (see Just et al. 1982, 85). As it turns out, another way to obtain CV and EV measures is to consider the Hicksian (rather than Marshallian) demand curve. The Hicksian demand curve is "a relationship giving quantities demanded at various prices when utility is held constant by varying income (alone)" Just et al. 1982, 87). The compensating variation equals the area to the left of the Hicksian demand curve going through the bundle in *A*; the equivalent variation equals the area to the left of the Hicksian demand curve going through the bundle in *B* (Just et al. 1982, 89).

These measures have certain advantages over consumer surplus measures, and therefore are used in many contexts to assess changes in welfare. As Charles Blackorby and David Donaldson (1990) write:

> In cost-benefit analysis and other exercises in applied welfare economics, *aggregate willingness-to-pay* – the simple sum of Hicksian compensating variations, is often used as a test. A positive sum is taken as evidence of a social improvement or an increase in economic efficiency (Blackorby and Donaldson 1990, 472, italics in original).

The value of the compensating or equivalent variation can be assessed by using econometric techniques to estimate the shape of the Hicksian demand function, or by simply asking people e.g.

about their willingness-to-pay. Another manner is to use the consumer surplus measure (discussed in the previous section) as an approximation (Just et al. 1982, 97-113). Willing (1976) has developed an expression for error bounds on consumer surplus as an approximation of CV or EV, and argues that the error is often relatively small (cf. Just et al. 1982, 114). Just et al. conclude: "These empirical embellishments of Hicks' conceptual results thus provide a sound foundation for consumer welfare measurement" (Just et al. 1982, 114).

1.2.3 The social indicator movement

The social indicator movement – so identified by Otis Dudley Duncan (1969, 1) – arose as a reaction to the widespread use of economic measures of well-being (Carley 1981, 1).[8] Members of this movement admitted that economic measures had some appealing features. The most important advantage, according to Angus Campbell, is that economic measures are "easy to count," and that their "units are equal and interchangeable" (Campbell 1976, 117). However, he added: "None of us doubts that economic data have admirable qualities; the question is, How well do they represent the quality of national life? How valid are they as measures of the goodness of life in this country?" (Campbell 1976, 117). He goes on to argue that, in the past, rapid economic development has not typically been associated with a comparable increase in quality of life.

Some proponents of the social indicator movement argued that economists had lost track of the distinction between means and ends.[9] Thus, the authors of the *Human Development Report*, published by the United Nations Development Programme (UNDP), wrote: "Caught up with the

[8] Robert J. Rossi and Kevin J. Gilmartin (1980) trace the history of the social indicator movement back to William Ogburn's work at the University of Chicago during the 1920's and 30's (Ross and Gilmartin 1980, 1).

[9] Others saw the social indicator movement as a move away from measures of well-being itself, and toward measures of the basic requisites of well-being. See Rescher (1972); cf. below.

rise and fall of national incomes, economists often lost sight of the real end of development – people's well-being. Economic growth is merely a means – albeit an important one – for achieving this end" (UNDP 2004, 127). Similarly, Frank Andrews (1989) argues that the social indicator movement "involved an attempt to focus directly on 'output' indicators – i.e. indicators that show how well off people actually are – in addition to the more traditional 'input' indicators" (Andrews 1989, 401). Notice the emphasis on well-being as the true goal of development. The UNDP adds that "human outcomes do not depend on economic growth and levels of national income alone. They also depend on how these resources are used – whether for developing weapons or producing food, building palaces or providing clean water" (UNDP 2004, 127).

By contrast, the social indicator movement sought to find "a broader and more sensitive set of measures that will provide a fuller description of people's lives" (Campbell 1976, 118). As Robert J. Rossi and Kevin J. Gilmartin (1980) put it:

> Interest in developing social indicators has been motivated by the desire to create a system of social accounts – analogous to the existing system of national economic accounts – that could be used to assess periodically the levels of social well-being, social effects of economic conditions, and the success of governmental programs (Rossi and Gilmartin 1980, 15).

Indeed, some have proposed the development of a comprehensive index of Gross National Welfare, analogous to the Gross National Product (Rossi and Gilmartin 1980, 27). Anyway, there is little doubt that many members of the social indicator movement wanted their collection of statistics to play the very same role as traditional economic accounts had come to play. In particular, the set of social indicators were supposed to provide a superior guide to public policy than economic indicators alone.

In practical terms, this movement encouraged the collection of data on life expectancy, quality of food and water, access to adequate medical care, level of education, quality of housing,

and so on, in order to attain a better measure of the people's well-being or quality of life. As Campbell notes: "It is reasonably argued that as the level of education rises, the adequacy of medical care improves, the amount of substandard housing is reduced, and the purity of the air and water is increased, the quality of life is therewith enhanced" (Campbell 1976, 118). As these indicators "do not depend on the individual's description of his own life," Campbell concludes that they "may be called *objective* indicators" (Campbell 1976, 118, italics in original).

Quite arguably, the most famous outgrowth of the social indicator movement is the Human Development Index. Since 1990, the Index has been published annually in the *Human Development Report* (most recently, UNDP 2004). The Human Development Index is a comprehensive index intended to offer "a powerful alternative to income as a summary measure of human well-being" (UNDP 2004, 137). As the UNDP write:

> The human development index (HDI) focuses on three measurable dimensions of human development: living a long and health life, being educated and having a decent standard of living. Thus it combines measures of life expectancy, school enrolment, literacy and income to allow a broader view of a country's development than does income alone (UNDP 2004, 128).

The authors of the UNDP report are aware of the fact that the choice of indicators is a non-trivial task. They write: "The range of capabilities that individuals can have, and the choices that can help to expand them, are potentially infinite and vary by individual" (UNDP 2004, 127). The authors claim to have used two criteria in the selection process: "First, these capabilities must be universally valued. Second, they must be basic to life, in the sense that their absence would foreclose many other choices" (UNDP 2004, 127).

In computing the index, the UNDP includes three different statistics: life expectancy at birth, the adult literacy rate along with a measure of school enrolment ratios, and the logarithm of GDP per capita (PPP USD) (UNDP 2004, 259). The HDI is computed as the simple average of

three dimension indices. After a maximum value and a minimum value for each underlying indicator have been determined, each dimension index is computed as follows:

$$\text{Dimension index} = \frac{\text{actual value} - \text{minimum value}}{\text{maximum value} - \text{minimum value}}$$

Thus, each dimension index, and the HDI as a whole, will be a number between zero and one (UNDP 2004, 259).

1.3 SUBJECTIVE MEASURES

In this section, I offer a preliminary discussion of subjective measures of well-being and of how they differ from the other kinds of measure discussed in the previous section. Because these issues will be discussed at greater length in the body of the dissertation, this section will be fairly brief. Nevertheless, I want to begin by outlining when subjective measures came from and the role that their proponents think that the measures can play. This should give us a better idea of what is at stake in the choice between different measures of well-being.

In spite of the fact that subjective measures are often described as a recent phenomenon, their history goes back a long time. As I show in chapter 2.0, that history can be traced back to the 1920's and 30's, when they were used in both theoretical and applied work in the domains of marital success and educational psychology. During this era, the measures were seen as tools that could be used to make sure that marriages lead to babies, and that education leads to happiness. In many cases, the work was seen as a straightforward application of classical utilitarianism.

As a large-scale measure of social well-being or welfare, however, the subjective measures did not gain currency until in the 1960's. At that time, they seem to have appeared as an unintended byproduct of the social indicator movement. This movement had turned against the traditional economic approach to welfare measurement (which had been in place at least since Pigou, the father

of welfare economics, declared that the national divided could be used as a measure of welfare). Some proponents of the social indicator movement objected that income and other economic measures were hopelessly indirect measures of welfare, and that they did not in fact adequately reflect real levels of well-being. Thus, the critics argued for the use of a richer set of measures supposedly more reflective of actual welfare, including access to health care, apartment size, and the number of telephones. Yet, once that step had been taken, it is hard not to admit that so-called objective indicators remain an imperfect measure of actual welfare. Thus, measures of happiness and satisfaction were defended as "direct" measures of well-being.

As already mentioned, subjective measures of well-being are designed to take into account people's subjectively experienced mental states. In the typical case, subjects are asked to fill in questionnaires with questions of the form "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?" In order to find out what levels of happiness are associated with phenomena like unemployment, the researchers then compare the average happiness score of unemployed subjects with that of employed ones. More recent studies tend to ask questions not only about happiness, but also about satisfaction, e.g. "How satisfied are you with your life these days?" Andrews and Withey (1976) noted that satisfaction ratings do not correlate very strongly with happiness ratings, and concluded that satisfaction and happiness are "separable constructs" (cf. Andrews and Withey 1976; cf. Diener et al. 1999, 277).

A somewhat different approach has been developed by Daniel Kahneman and co-authors (e.g. Kahneman et al. 1999). Kahneman prompts his subjects every so often – e.g. with the use of palm pilots – to judge the "quality of their momentary experience" along the "good/bad dimension" (Kahneman 1999, 7). He writes: "Two separate assumptions are involved: that the brain continuously constructs an affective or hedonic commentary on the current state of affairs, and that

this commentary is adequately summarized by a single value" (Kahneman 1999, 7). At every point in time, the brain rates the qualities of experience in a manner that can be represented on a single numerical scale (and which, furthermore, is accessible to the agent). In Kahneman's terms, then, well-being is a matter of positive hedonic tone. Kahneman also introduces another important distinction, that between subjective and objective well-being (or happiness). He writes: "We distinguish two notions of happiness, or well-being (the two terms are used interchangeably in this chapter). *Subjective happiness* is assessed by asking respondents to state how happy they are. *Objective happiness* is derived from a record of instant utility over the relevant period" (Kahneman 1999, 5).[10] More specifically, the objective happiness during some period of time is computed by taking the time integral of the subjective happiness. Kahneman and co-authors have since developed other measures, though Kahneman et al. (2004) insist: "Experience sampling is the gold standard" (Kahneman et al. 2004, 1777).

Relying on these measures, psychologists claim to have confirmed the existence of several important phenomena. For one thing, psychologists have suggested that there is only a weak relation between income and happiness.[11] As Ed Diener and Robert Biswas-Diener write, "for middle and upper-income people in economically developed nations, acquiring more income is not likely to strongly enhance SWB. Indeed, some studies find that rising wages predict less well-being" (Diener and Biswas-Diener 2001, 161). It also appears that even rapid economic growth is not associated with measurable increases in subjective well-being (Diener and Biswas-Diener 2001, 139).

[10] The terminology here is unfortunate. Better terms would have been "momentary" (or "instant") and "overall" well-being, or some such.

[11] Results such as this one emphasize suggest that different measures will give rather different answers to questions about the determinants and distribution of well-being. Thus, for both practical and scientific purposes the choice of measure may matter a great deal.

Meanwhile, people who live in rich countries are on the average happier than people living in poor countries (Diener and Biswas-Diener 2001, 136). These phenomena – and others like them – are often explained by reference point phenomena (Argyle 1999), the process of adaptation (Frederick and Loewenstein 1999), and misprediction (Loewenstein and Schkade 1999; cf. Loewenstein and Angner 2003). Interestingly, many of these ideas are not novel; indeed, some of them can be found already in the work of Adam Smith (cf. Ashraf et al. manuscript) The psychological research, however, have emphasized both the wide variety of conditions under which those phenomena occur, and just how strong they can be.

The reason why psychologists care so much about subjective measures of well-being is that they believe well-being – as they understand it – has a privileged normative status. Some go so far as arguing that public policy should be designed so as to maximize well-being. This appears to be the view of Richard Layard, in his recent book *Happiness:* "[Bentham] proposed that all laws and all actions should aim at producing the greatest possible happiness.... I believe that Bentham's idea was right and that we should fearlessly adopt it and apply it to our lives" (Layard 2005, 111-112). Later on, Layard specifies that the principle holds not only in private life but also for public policy (Layard 2005, 115). Oswald (1997) makes a similar point when he writes: "The relevance of economic performance is that it may be a means to an end. That end is ... the enrichment of mankind's feeling of well-being. Economic things matter only in so far as they make people happier" (Oswald 1997, 1815). Andrews and Robinson defend their focus on well-being in the following way:

> Subjective well-being is important as a psychological summing up of the quality of an individual's life in society. Several social psychological concepts tap aspects of the quality of life indirectly, such as self-esteem, depression, locus of control, or alienation, but only life satisfaction and happiness have a "bottom-line" finality in terms of consequences for the individual (Andrews and Robinson 1991, 61).

A similar point of view is expressed in the following quote:

> 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 (Campbell 1976, 119).

Kahneman writes: "In the present framework ... it is objective happiness that matters. Policies that improve the frequencies of good experiences and reduce the incidence of bad ones should be pursued" (Kahneman 1999, 15).[12] Finally, Diener and Seligman assert: "Our thesis is that well-being should become a primary focus of policymakers, and that its rigorous measurement is a primary policy imperative.... [We] propose that well-being ought to be the ultimate goal around which economic, health, and social policies are built" (Diener and Seligman 2004, 1-2). These quotes raise a great number of questions; for now, it is sufficient to note that psychologists assume that subjective well-being – as they understand it – has an important normative status.

In light of these ambitions, it is unsurprising that psychologists should have many suggestions about how to make people happier and the world better. Books like *The Loss of Happiness in Market Economies* (Lane 2000), *The High Price of Materialism* (Kasser 2002), and *The Progress Paradox* (Easterbrook 2003) rely on research into happiness, satisfaction, and the like, to identify how public policy and our personal lives and can be reoriented so as to be more effective at promoting subjective well-being. Many of the proposals derive from the finding that the marginal happiness of money for the affluent appears to be close to zero. Thus, it is sometimes suggested that it is justified to redistribute resources from the financially well off (who are not expected to derive much happiness from those last dollars) to the poor and unemployed (who are). Similarly, it is often suggested that it is justified to work less (since the additional income is not likely to promote

[12] Note that "objective happiness," as Kahneman uses the term, is derived from ratings of subjective happiness or well-being. It does not refer to objective indicators, in Campbell's terms.

happiness much), and spend more time with family and friends (since that is likely to promote happiness).

The enthusiasm for the usefulness of the new measures is almost boundless. According to Kahneman et al. (2004), measures like those they discuss are

> ... potentially useful to medical researchers for assessing the burden of different illnesses (1) and the health consequences of stress (2); to epidemiologists interested in social and environmental stressors (3); to economists and policy researchers for evaluating policies and for valuing non-market activities (4, 5); and to anyone who wishes to measure the well-being of society (Kahneman et al. 2004, 1776).

For the latter purpose, Kahneman et al. (2004) have suggested that we establish "national well-being accounts" (NWBA) analogous to national accounts regarding production and so on. Similarly, Diener and Seligman (2004) have argued that what we need is a "set of national indicators of well-being" (Diener and Seligman 2004, 21). They write:

> The most important contribution of a national system of well-being indicators would be that they could focus the attention of policymakers and the public specifically on well-being, and not simply on the production of goods and services; one of the main benefits of well-being measures is that they add a valuable perspective beyond a cost-benefit market analysis in evaluating societal structures and interventions (Diener and Seligman 2004, 21).

Diener and Seligman argue that policymakers already care about well-being, but that their attempts to promote it are based on "mere guesses and romantic sentiments" (Diener and Seligman 2004, 21). The hope is, obviously, that national indicators of well-being will allow policymakers to pursue well-being in a scientifically informed manner.

Coincidentally or not, some of the psychologists who so strongly advocate the use of subjective measures for all sorts of purposes also stand to gain economically from the popularization of these measures. The Gallup organization says it spends "huge" amounts of money on developing these measures, presumably on the assumption that they will generate commensurate revenues in the

future (Daneshkhu 2005). In some of their recent publications, both Diener and Kahneman are identified as affiliated with the Gallup Organization (Diener and Seligman 2004; Kahneman et al. 2004).

1.4 THE POINT OF THE PROJECT

There are, in my view, a number of reasons why subjective measures deserve our attention. From a more scientific point of view, a philosophical investigation into the conceptual foundations of measures of well-being can serve several functions. Clarity about conceptual foundations can (i) make it easier to properly assess the different measures, (ii) clarify the significance of the empirical findings, (iii) help identify whether (or under what conditions) the measures can serve as a basis for policy, and (iv) remove obstacles to fruitful communication and cooperation across disciplinary boundaries. Below, we will see several examples of problems that can occur in the absence of such clarity.

From a more philosophical perspective, there are some clear benefits as well. To begin with, the literature by economists and psychologists raises interesting issues about, and sometimes offers tentative answers to, purely philosophical questions e.g. about the nature of well-being. Thus, Sumner (1996) draws heavily on the psychological research in defending an account of well-being as life satisfaction. Also, many philosophical arguments in ethics and political theory rest in part on empirical premises. This, I believe, is what John Rawls was alluding to when he wrote that "the fundamental principles of justice quite properly depend upon the natural facts about men in society" (Rawls 1999, 137). Surely, insofar as the principles of justice depend on empirical facts about the determinants and distribution of well-being, the psychological and economic literature (properly interpreted) is directly relevant to the theory of justice. Similarly, Nicholas Rescher (1978) – relying

on the authority of Immanuel Kant – notes: "In social philosophy the actualities of empirical circumstances must ever predominate" (Rescher 1972, ix). The point is that many arguments in ethics, and in social and political philosophy, proceed from, and depend on, empirical facts.

Moreover, it is sometimes argued that accounts of well-being can be judged in part on the grounds of whether they permit the development of valid measures of well-being. This idea is explicit in a number of prominent contemporary philosophers, including James Griffin (1986, 1) and Christine M. Korsgaard (1993, 54). As we will see below, it also appears to be implicit in economic literature on welfare measurement, from the 1930's, or earlier, to the present. If this is correct, then the success (or lack thereof) of attempts to measure e.g. subjective well-being is eminently relevant to the adequacy of their account of well-being. If so, the results of this investigation have implications for the theory of well-being and all ethical theories that give the concept some role.

Finally, the issue about what measure of well-being should be used in the designed and evaluation of public policy is of obvious political and practical importance. It is widely (though not universally) agreed that public policy should be designed (at least in part) so as to promote the well-being of the population (see e.g. Ahlheim 1998, 484). It would not matter which measure we used if they tended to give the same answer to the question about who is well off, and about what conditions tend to make people better off. Interestingly, however, judging by available empirical research the measures do not agree in this way. While it may seem obvious that well-being should be strictly increasing in wealth, for example, this appears to be true only with modification. Insofar as policy should be designed to promote well-being at all, the choice of measure will have real consequences for the policies that we favor. The example of Russian post-Soviet reforms (see Angner in progress) will illustrate this point.

In spite of the attention that psychologists have given to subjective measures of well-being, so far they have received relatively little attention from philosophers. To my knowledge, the first

philosopher who dedicated serious attention to this literature was Rescher, in his books *Welfare* (1972)[13] and *Unpopular Essays on Technological Progress* (1980).[14] More recent philosophers who have discussed the psychological findings include L. W. Sumner (1996) and Daniel M. Haybron (2000; 2003). However, Sumner and Haybron are primarily interested in how to develop accounts of well-being and happiness, respectively, and therefore leave many (in my view interesting) questions unanswered. Specifically, neither addresses the issue of well-being measurement in any detail. The fact that subjective measures of well-being have received so little attention, I believe, makes it particularly fertile ground for a historically informed philosophical analysis.

1.5 DISCLAIMERS

Before closing the introduction, I should say a few words on what I do not presume to do. I do not believe that I can identify every aspect of the philosophical foundations of subjective measures, or all the advantages and disadvantages associated with them. Surely, this would be an insurmountable task. I only aspire to discuss some aspects – albeit, in my view, important ones – of the subjective measures, especially as they compare to economic ones. Moreover, it is not my goal to defend any particular account of well-being. It is true that the discussion may bear on the question of which account of well-being is the most plausible, but my immediate goal is simply to identify the account of well-being that underlies some prominent measures of well-being, and identifying and discussing advantages and disadvantages associated with these measures. Finally, it is not part of my project to argue for the claim that well-being should be promoted. Here, I want to prepare the ground for addressing the following issue: Insofar as we are concerned with promoting people's well-being, and

[13] See especially chapter 3, 'Social Welfare and Personal Happiness,' (Rescher 1972, 36-59).

[14] See especially chapter 1, 'Technological Progress and Human Happiness,' (Rescher 1980, 3-22).

insofar as we need a measure of it, what considerations speak for and against subjective measures (as compared to traditional economic ones)?

2.0 THE EVOLUTION OF EUPATHICS: THE HISTORICAL ROOTS OF SUBJECTIVE MEASURES OF WELL-BEING

Abstract. In this chapter, I explore the history of subjective measures of well-being. While it is often suggested that these measures are a fairly modern invention, I argue that they have a long and rich history. Subjective measures appear to have emerged as a result of studies into educational psychology, marital success studies, and personality psychology in the 1920' and 30's, and evolved under the influence of the epidemiology of mental health, gerontology, and the social indicator movement in the 1960's and 70's. The story confirms the main conclusions drawn by Theodore Porter (1995) in his general discussion of measurement in the social and behavioral sciences: these measures emerged in applied rather than theoretical branches of social science, and they did so not as a result of physics envy, but rather as a result of a moral impulse to improve society; quantification was intended to make up for perceived deficiencies in unaided human judgment; and radical disagreements about the nature of well-being did not impede efforts to measure it.

2.1 INTRODUCTION

It is often suggested that subjective measures of well-being constitute a relatively novel development. Bruno S. Frey and Alois Stutzer (2000), for example, write: "Recently, great progress has been achieved in economics: happiness has been seriously measured, and many of its determinants have been identified" (Frey and Stutzer 2000, 145). In a book published in 2002, the same authors point out that economists traditionally have given little attention to questions of happiness, and add: "In the past few years the situation has changed: A number of economists see an advantage in measuring subjective well-being as expressed by individuals themselves" (Frey and

Stutzer 2002, vii). Daniel Kahneman et al. (2004) add: "Economists have traditionally eschewed direct measures of well-being" (Kahneman et al. 2004, 429).

The impression that subjective measures emerged fairly recently is partially due to the fact that they are often linked to the positive psychology movement. This movement, which emerged in the 1990's, was motivated by the belief that traditional psychology had spent an inordinate amount of time examining pathology.[15] Martin E. Seligman writes: "For the last half century psychology has been consumed with a single topic only—mental illness" (Seligman 2002, ix). According to the positive psychologists, traditional psychology overlooked positive emotions because they were seen as derivative from, or less authentic than, negative emotions, and therefore less worthy of study. In contrast, positive psychology proceeds from the assumption that positive emotions are no more derivative (and no less authentic than) negative ones, and therefore worthy of attention in their own right. It goes without saying that the study of subjective well-being is an integral and important part of positive psychology. Because of their implicit or explicit association with positive psychology, it may seem that subjective measures of well-being constitute a relatively novel development too.

The goal of this chapter is to explore the historical roots of subjective measures of well-being. My thesis is that such measures, far from being a novel invention, have a long and rich history. I claim that they can be traced back at least to the 1920's and early 1930's. They emerged when psychologists and psychiatrists interested in educational psychology and marital happiness began to answer empirical questions about happiness, satisfaction, and well-being in a systematic fashion. This development was a consequence of the rise of personality psychology after World War I, and was further shaped by the epidemiology of mental health, gerontology, and the social

[15] See Gillham and Seligman (1999), the special January 2000 issue of *American Psychologist*, especially Seligman and Csikszentmihalyi (2000), and the massive *Handbook of Positive Psychology* (Snyder and Lopez 2002). This paragraph draws primarily on Seligman and Csikszentmihalyi (2000, 5-9) and on the preface to Seligman (2002).

indicator movement in the 1960's and 70's. Meanwhile, the researchers in this tradition had rather different purposes, used a wide variety of concepts, and invented a number of tools to measure it.

The story, I find, confirms the general outlines of Theodore Porter's account of measurement in the social and behavioral sciences. His account is most clearly developed in the book *Trust in Numbers: The pursuit of objectivity in science and public life* (1995), and has been discussed and further clarified in a number of forums, e.g. in the collection *The Age of Economic Measurement* (Klein and Morgan 2001) and the special issue of *Studies in History and Philosophy of Science*, Vol. 32, No. 4 (2001), edited by E. P. Hamm and Alan W. Richardson. Consistent with Porter's conclusions, I find that subjective measures of subjective well-being emerged in applied rather than theoretical branches of social science, where they were developed not as a result of physics envy, but of a moral impulse to improve society; that quantification was intended to make up for perceived deficiencies in unaided human judgment; and that radical disagreements about the nature of well-being did not impede efforts to measure it.

In tracing the history, I take as my starting point the references in the review article by Ed Diener, Eunkook M. Suh, Richard E. Lucas, and Heidi L. Smith (1999), and follow the paper trail backwards. For the early studies, I have relied primarily on the many references in the review article by Warner Wilson (1967), and the sources they quote. In discussing the scientific studies, I have focused on what motivated them in the first place, what the underlying notion of well-being – and happiness, satisfaction, mental health, and so on – was, how the degree of happiness (satisfaction, mental health) was measured, and what phenomena the authors claim to have discovered.

2.2 THE HISTORY OF SUBJECTIVE MEASURES

In this section, I discuss the early efforts to develop measures of happiness, satisfaction, and well-being.[16] I have grouped each study by the intellectual domain in which it appeared. This allows me to emphasize the broad movements (or traditions) that generated an interest in happiness and the like, and in developing accurate methods to measure it. I do not mean to suggest that the traditions were entirely independent, however. As I try to indicate in the text, there was a fair amount of cross-pollination between different fields.

2.2.1 Marital success

Two of the earliest studies on happiness are penologist and social worker Katharine Bement Davis' *Factors in the Sex Life of Twenty-Two Hundred Women* (1929) and psychiatrist G. V. Hamilton's *A Research in Marriage* (1929). Davis, who received her Ph.D. in economics, was interested in gathering "adequate data as to both the physical and mental facts of the sex life of the normal individual" (Davis 1929, ix). She also wanted to explore correlations between facts about the women's sex life and other aspects of their lives. Thus, one of the questions given to the roughly 2200 subjects was the following: "Do you consider your life on the whole (a) happy, satisfactory, successful; (b) unhappy, unsatisfactory, unsuccessful? In each case why?" (Davis 1929, 89).

Hamilton's study was similar in spirit, except that his subjects were married couples. He gives a window into what motivated these studies when he writes:

[16] It is always difficult to know where to begin a history of any scientific development. I could, for instance, have started by discussing the work of A. Wohlgemuth (1919) and J. C. Flügel (1925), whose research program had many features in common with later research on subjective well-being. However, from what my research has shown, these studies were never as influential as the ones I mention here, and have therefore been omitted.

> My standpoint is that of the psychiatrist who believes that subjective phenomena, as these are experienced by the persons who report their occurrence, do not need to be translated into anything else in order to be dealt with as objectively as we deal with all other biological phenomena (Hamilton 1929, xi).

Specifically, Hamilton wants to know how satisfied his subjects are with their marriages (Hamilton 1929, chapter 2.0). To this end, he used responses to a set of questions about satisfaction to compute a "satisfaction-grade," which he also refers to as an "index of spousal satisfaction/dissatisfaction" (Hamilton 1929, 78-79). Hamilton took the satisfaction-grade to signal what he called "marital success" (Hamilton 1929, 8). Anyway, Hamilton – and, presumably, Davis – believes that such subjective phenomena are worthy of study in their own right, and that they can be scientifically studied using psychological means.

Lewis M. Terman, drawing on Davis (1929) and Hamilton (1929) continued to explore marital success. In a book called *Psychological Factors in Marital Happiness* (1938), he describes his project in the following way: "We have selected as the theme of our study that aspect of the successful marriage which may be designated as marital happiness, and we wish to ascertain, if possible, what psychological factors are demonstrably associated with this state" (Terman 1938, 2). Terman immediately goes on to say that he is not committed to the view that "personal happiness is the only 'proper' goal of marriage," but then adds: "It is of the very nature of happiness that, other things being equal, it should be preferred to its opposite" (Terman 1938, 2).

Unfortunately, Terman does not say much about the meaning of "happiness." In order to avoid "philosophical connotations," he says, he "preferred to apply the term in a sense familiar to everyone," whatever that sense may be (Terman 1938, 3). There is some discussion about the scale on which happiness can be measured, however. In Terman's words:

> Happiness cannot be measured as distance is measured in terms of equal units on an unambiguous scale. Degrees of happiness are nevertheless very real, and the use of a numerical index is justified if the subjects it rates high are in fact definitely more happy than the subjects it rates low (Terman 1938, 4).

In order to develop a means to predict marital happiness and unhappiness, Terman and his collaborators distributed questionnaires to 792 couples from "middle and upper-middle classes of urban and semiurban Californians" (Terman 1938, 13). They constructed a *marital happiness score* (also referred to as an *index of satisfaction*) on the basis of "(1) subjective ratings of the happiness of the marriage; and (2) factual information on husband-wife agreement or disagreement about various matters," in which the subjective rating "was allowed a heavy weighting" (Terman 1938, 3; cf. p. 367). For the subjective rating, respondents were given the question "Everything considered, how happy has your marriage been?" (Terman 1938, 440). They were then asked: "Draw a circle around 1, 2, 3, 4, 5, 6, or 7," where the options were "1=Extraordinarily happy," "2=Decidedly more happy than the average," "3=Somewhat more happy than the average," and so on (Terman 1938, 440). Note that the question strictly speaking concerns the happiness of the marriage, rather than that of either individual in it.

On the basis of his results, Terman writes that he could develop a picture of the "happy and unhappy temperaments" (Terman 1938, 369). He writes:

> For example, it is especially characteristic of unhappy subjects to be touchy or grouchy; to lose their tempers easily; to fight to get their own way; to be critical of others; to be careless of others' feelings; to chafe under discipline or to rebel against orders; to show any dislike that they may happen to feel; to be easily affected by praise or blame; to lack self-confidence; [and so on] (Terman 1938, 369).

The background factors most strongly correlated with happiness, in Terman's study, were "Superior happiness of parents," "Childhood happiness," and "Lack of conflict with mother" (Terman 1938, 372). Terman also examines the relationship between happiness and various "sex factors" and

concludes that the influence of the sex factors is probably no greater than that of background and personality factors combined (Terman 1938, 375-376).

As we have seen, Terman offers few clues about what he has in mind by "happiness" and "satisfaction," though he does appear to use the terms interchangeably. Unlike many other authors, Terman creates a happiness score by combining the subjective rating with other "more objective" judgments, but from the weighting scheme it is clear that he considers the subjective rating the most important item. He asks subjects to consider their happiness during the entire marriage, rather than current happiness, and he offers subjects seven alternative answers.

A decade after Davis (1929), Ernest Burgess and Cottrell offer a study called *Predicting Success or Failure in Marriage* (1939), drawing above all on Terman (1938). They write that marital adjustment and incompatibility has become a social problem and therefore of public concern (Burgess and Cottrell 1939, 1). Defining their terms, the authors write:

> A well-adjusted marriage from the point of view of this study may then be defined as a marriage in which the attitudes and acts of each of the partners produce an environment which is favorable to the functioning of the personality of each, particularly in the sphere of primary relationships (Burgess and Cottrell 1939, 10).

Their study has three goals: to define "marriage adjustment," to identify what factors are associated with marital success or failure, and to determine whether it is possible to predict ahead of time what marriages will lead to happiness and which will lead to unhappiness (Burgess and Cottrell 1939, 15). In practice, the authors use happiness as the criterion by which they judge both adjustment and success (Burgess and Cottrell 1939, 30). Instead of "success" and "happiness" the authors also sometimes talk about "satisfaction" (cf. Burgess and Cottrell 1939, 45).

When it comes to the definition of "happiness," Burgess and Cottrell rely on *Webster's New International Dictionary*, which defines the term as "a state of well-being characterized by relative permanence, by dominantly agreeable emotion ranging in value from mere contentment to positive

felicity, and by a natural desire for its continuation" (Burgess and Cottrell 1939, 31). As the authors note, the dictionary definition emphasizes the subjective nature of the concept (Burgess and Cottrell 1939, 31). In their study, 526 couples were given the following question: "Appraisal of marriage: very unhappy.....; unhappy.....; average.....; happy.....; very happy....." and asked to check the relevant box (Burgess and Cottrell 1939, 422).[17] They made no attempt at explaining the meaning of "happiness" to their subjects, on the assumption that people in general would understand the notion in accordance with the dictionary definition (Burgess and Cottrell 1939, 31).

On the basis of their results, Burgess and Cottrell conclude that the happiness ratings are both reliable and stable. For each marriage, they compared the rating of the husband to the rating of the wife – the questionnaires were supposed to be filled in independently – and found that only 3.6 percent differed by more than one scale step (Burgess and Cottrell 1939, 38). The authors also compared the ratings of individual spouses with that of a knowledgeable outsider, and found that only 8.8 percent differed by more than one scale step (Burgess and Cottrell 1939, 41). Burgess and Cottrell also inferred that the happiness ratings were stable, since they found little change in ratings after a period of eight to 24 months (Burgess and Cottrell 1939, 43-44).

In sum, Davis (1929), Hamilton (1929) and their followers were interested in marital adjustment, marital success, and marital happiness, and how to predict it on the basis of personality, background and sexual factors. Interestingly, at least some of the researchers in this tradition were often more interested in the happiness of the marriage, that in that of the people in it. One possibility is that their concern with "marital success" was motivated not primarily by a concern to make people happy, but by a desire to insure a sufficiently high population growth rate.[18] Either way,

[17] Again, the question strictly speaking concerns the happiness of the marriage, as opposed to that of the husband and wife.

[18] I owe this suggestion to Prof. Mark Perlman.

it is clear that they took phenomena such as happiness to be worthy of study, that they thought it possible to develop scientific measures of marital happiness, and that they though it worth promoting.

2.2.2 Educational psychology

Perhaps the most prominent early study of subjective well-being is Goodwin Watson's 'Happiness Among Adult Students of Education' (1930). Watson, a professor of education at Columbia,[19] introduces his topic in the following way:

> No human quest may claim a larger following that that for happiness and satisfaction in life. Even the highest ethics tends to justify itself by its contribution to human happiness.... Certainly any educational program in modern times is likely to be justified only in terms of its direct or indirect contribution to human happiness. It becomes, therefore, extraordinary almost beyond belief that so few attempts have been made to apply the techniques of psychological study to the understanding of happiness (Watson 1930, 79).

By "the highest ethics," presumably, Watson refers to the utilitarians' maxim of the greatest happiness for the greatest number, and he is objecting to the fact that psychologists have not set out to explore, in a systematic, way, how to achieve greater happiness. Watson sets out to make up for the deficiency. Unfortunately, he does not specify what he means by "happiness" (or "satisfaction"), though the introductory remark suggests that he takes himself to be using the same concept as some utilitarian. He adds: "What is studied might, in the strictest sense be termed not happiness, but self-estimates of happiness" (Watson 1930, 79). Yet, the fact that he from that point on consistently talks about happiness *tout court* suggests that in Watson's mind the two are closely connected.

Watson distributed anonymous questionnaires to 388 graduate students of education, and used a variety of methods to elicit their degree of happiness (Watson 1930, 79). First, he asked:

[19] It goes without saying that Goodwin Watson is not John B. Watson, author of *Behaviorism* (1924).

"Comparing yourself with other persons of the same age and sex how do you feel you should rate your own general happiness?" (Watson 1930, 79). The subjects were then given a horizontal line, with captions that read (from left to right) "Most miserable of all," "About three-fourths of the population happier than you are," "The average person of your own age and sex," and so on (Watson 1930, 80). Subjects were asked to make a short vertical mark at the point where they thought they belonged, taking into account their "average state over several months" (Watson 1930, 80). The subjects were also to place a circle at the point where they thought their friends would rate them (Watson 1930, 80).

Second, subjects were offered ten descriptions of one or two sentences, and asked: "Among the following descriptions, arranged in miscellaneous order, choose the one which comes nearest to fitting you" (Watson 1930, 80). The descriptions included the following: "Cheerful, gay spirits most of the time. Occasionally bothered by something but can usually laugh it off," "Ups and downs, now happy about things, now depressed. About balanced in the long run," and "Life often seems so worthless that there is little to keep one going. Nothing matters very much, there has been so much of hurt that laughter would be empty mockery" (Watson 1930, 81).

Third, subjects were given a blank space and asked: "Now write in a sentence or two, something like those above, which you believe will most truly describe your own general happiness in life" (Watson 1930, 81). Fourth, subjects were given a list of fifty properties – "Enthusiastic," "Troubled," "Annoyed," and so on – and asked: "Check every term which you believe could fairly be applied to yourself in your prevalent attitudes" (Watson 1930, 81). Half of the descriptions were positive, half negative (Watson 1930, 81). Finally, using a graphic rating device as in the first question, subjects were asked about their happiness in different areas, viz. "health, vocation, love and marriage, friends, hobby interests and religion," and in different stages of life, viz. "early childhood ..., later childhood ..., high school period, and later adolescence" (Watson 1930, 82).

On the basis of subjects' responses, Watson computed an aggregate "happiness score" (Watson 1930, 82). He writes: "Each graphic rating scale was scored by a scale of units ranging from 0 at the most unhappy extreme to 100 at the happiest" (Watson 1930, 82). The descriptions in the second part of the questionnaire were assigned values from 1 to 10 and multiplied by 7 for greater weight (Watson 1930, 82). The response in the third (subjective) part was scored on a scale from 1 to 10 by three judges, and the median rating was multiplied by 6. To score the fourth part, Watson counted the number of positive traits mentioned and subtracted the number of unhappy ones. For the areas of experience, Watson included only the average in the aggregate measure (Watson 1930, 82).

To check for reliability, Watson used the responses to different items to create two different happiness scores. He noted that the correlation between the scores for men and women are as high as .83 and .85. As a result of these calculations, he concluded: "This indicates that the measure is reasonably consistent, throughout. It does not, of course, answer the other very interesting question but as yet unanswered question of the stability of this measure from day to day, and from week to week" (Watson 1930, 83). Watson also notes that the correlation coefficient between the subjects' own rating of their happiness on a graphical scale and the aggregate happiness score for men and women is .81 and .82 (Watson 1930, 86). Thus, he infers that for some purposes, the self-rating may suffice. He writes: "For the purpose of separating high and low groups such a simple indicator would probably be adequate" (Watson 1930, 86).

Examining subjects' responses, Watson noted that subjects in general "tended to think of themselves as above the average on each happiness item" (Watson 1930, 84). Moreover, those subjects who scored low to average in aggregate happiness "appeared to believe that they gave an impression of greater happiness than was justified," whereas those who scored high "felt no sense of masquerade" (Watson 1930, 87). Furthermore, Watson studied the correlation of aggregate

happiness and various traits of the subjects and living conditions. For intelligence, Watson had access to IQ scores; for the other traits and conditions he relied on information elicited from the subjects. Watson found that happiness was not predicted by intelligence (Watson 1930, 88), age, whether the subject came from a small town, physical defects, number of siblings, parents' demand for obedience, having parents older than 40, whether the mother had a career, parents' divorce or separation, school marks, extra-curricular activities, crushes, a wise sex education, or (perceived) popularity with the opposite sex, among other things (Watson 1930, 94-96). Perceived harmony between parents did predict happiness, however (Watson 1930, 94), as did among other things a preference to spend a few hours with the most delightful known companion of the same sex over one of the opposite sex, an ability to give a lecture to high school students about sex, an ability to administer a large group of people, an absence of strong fears, health during adolescence, an ability to fuse well, an absence of shyness or timidity, and being married (Watson 1930, 97).

Watson acknowledges a number of shortcomings of his study. First, he is aware of the fact that his sample is not representative of the general population (cf. Watson 1930, 88). Watson also at least implicitly acknowledges that conclusions about the causes of happiness cannot be drawn on the basis of correlational data alone. He does speculate about the causes of happiness, but ends by stating: "The proportion of error in these statements is uncertain, but considerable" (Watson 1930, 109).

In effect, Watson (1930) sets the tone for much of the research that followed. He is clearly interested, in the first place, in subjectively experienced happiness. He uses questionnaires as a means to explore the degree of happiness enjoyed by his subjects. Unlike the authors on marital happiness he uses a variety of tools, but finds that subjects' self-ratings are sufficiently highly correlated with the aggregate happiness measure to be used as a sole measure (at least for some purposes). Unlike later authors, Watson does not frame his project in terms of well-being or welfare.

Yet, he suggests that his notion of happiness is the same as that of the utilitarians, which makes it interesting from the point of view of the measurement of well-being. He talks about satisfaction as well as happiness (e.g. in the first sentence of his paper), and there is no evidence that he uses the terms in any other way than as synonyms.

The other early article in educational psychology about the determinants and distribution of happiness is 'Happiness self-estimates of young men,' by Randolph Sailer (1931). Sailer followed the approach taken by Watson, but he was also inspired by among others Hamilton (1929) (cf. Sailer 1931, 7). This fact indicates that there was cross-pollination at an early stage. Researchers in different fields were not unaware of each other's efforts, and they appear to have allowed their research design to be affected by what others had done and what they had learned. Distributing questionnaires to 500 young men across the U.S., Sailer found (among other things) a connection between happiness with religion and sociability (Sailer 1931, 98-99). Like Watson, Sailer concluded that for many purposes simpler measures of happiness could be substituted for more complex ones (Sailer 1931, 100).

In the mid-1930's a number of studies appeared, following Watson and Sailer, which explored the relationships between happiness scores and a number of variables. Thus, George W. Hartmann's 'Personality Traits Associated with Variations in Happiness' (1934), finds little correlation between personality traits and experienced happiness (Hartmann 1934, 211). Incidentally, Hartmann – who was a psychologist at the Pennsylvania State College – also uses the term *felicity* instead of happiness (cf. Hartmann 1934, 209). Similarly, it may be worth mentioning Percival M. Symonds' 'Happiness as Related to Problems and Interests' (1937), which concludes: "Happy and unhappy are remarkably alike in their problems and interests" (Symonds 1937, 293). Like Watson, Symonds was associated with the Teachers College at Columbia.

In passing, Hartmann (1934) makes favorable mention of Abraham Myerson's program of *eupathics* (Myerson 1918, 343). Myerson was a Clinical Director at the Taunton State Hospital in Massachusetts and a Harvard neuropathologist. His goal was to establish a field of Mental Hygiene that had "for its aim the *well being of the normal*" (Myerson 1918, 344, italics in original). He called the program the "more gracious sister" of eugenics, which "largely simmers down to a program for the elimination of the unfit" (Myerson 1918, 344). Myerson had great hopes for the field. He appears to have equated well-being with happiness, and believed that "Mood" was an important determinant of happiness. Furthermore, he claimed that mood "can be reached and elevated in a perfectly definite manner" (Myerson 1918, 344). He invested eupathics with a great deal of importance. After asking rhetorically what was left out of the field, he answered: "nothing of consequence" (Myerson 1918, 346). This suggests that, according to Myerson, well-being is the only thing that ultimately matters.

In this subsection, I have discussed authors whose works appears in the context of educational psychology. Watson's paper indicates that they may have been motivated at least in part by general utilitarian considerations, more specifically, by a desire to explore scientifically how one may go about promoting the greatest happiness for the greatest number. In particular, the research of Watson et al. appears motivated by a belief that education should aim to increase the happiness of the students. Their interest may also have been fueled by a frustration with the lack of systematic data on the basis of which to judge whether education in fact made students happier, and whether it was possible to improve the education so as to promote students' happiness, and if so how.

2.2.3 Personality psychology

Though the research program on subjective well-being, as we have seen, has its roots in research on marital happiness and educational psychology, the discussion also points strongly to the field of personality psychology. According to a modern definition: "Personality psychology is the scientific

study of the whole person. The goal of personality inquire is to provide a scientifically credible account of human individuality" (McAdams 2001, 11308). At the time when Watson and his followers were publishing on happiness, personality psychology was in the process of emerging as an independent subdiscipline of psychology. Beginning in the late nineteenth century, psychologists had made a serious effort to develop reliable methods for studying differences in personality traits across people and over time. In the words of David G. Winter and Nicole B. Barenbaum (1999), the "1921-1938 period was a time of intense research activity," by the end of which the field of personality psychology was officially established (Winter and Barenbaum 1999, 8-9). The turning point is marked by the appearance of the first authoritative textbook (Allport 1937) and a landmark study (Murray 1938) (cf. McAdams 2001, 11309).

From its beginnings, personality psychology was characterized by an emphasis on measurement and psychometrics, and by a desire to be useful to corporations and governments (Winter and Barenbaum 1999, 5). In McAdams' words:

> Building on the pioneering work of Francis Galton and Alfred Binet on mental testing and spurred by the mobilization of large military forces in World War I, psychologists began to invent self-report, multi-item tests to assess individual differences in personality functioning (McAdams 2001, 11309).

The hope was that a new exact science of personality would "furnish assistance to a corporate culture and a government suddenly confronted by dramatic changes and the need to 'manage' and control an American population that had suddenly become larger, more diverse, and 'difficult'" (Winter and Barenbaum 1999, 5).

The development of measures of happiness and satisfaction during the late 1920's and 1930's should be seen against this background. The research on happiness surveyed above was obviously driven by a desire to measure and account for traits or other characteristics that differ across individuals. It was also motivated by an ambition to be of use to governments (though not, as

far as I can tell, to corporations). The discussion on marital happiness suggests that the goal is also to help individuals make better decisions concerning their own life.

Seeing the measurement of happiness as an outgrowth of personality psychology helps answering Watson's question at the very beginning of the previous subsection (see section 2.2.2). Watson (1930) asked why it had been so long since the articulation of the utilitarian principle – the greatest happiness for the greatest number – for psychologists to explore systematically who is happy and why. Why did it take some 150 years before psychologists began to explore empirically how happiness is distributed, what factors predict it, and how one should go about promoting the greatest happiness for the greatest number, if one should be so inclined? The link with personality psychology goes a long way toward answering this question. It was not until after Galton, Binet, and the experiences of World War I that psychologists developed the confidence that they could reliably measure personality characteristics like happiness.

2.2.4 Synthesis and further development

Relatively quickly, there emerged a number of studies that drew equally on the results on marital happiness and educational psychology. For example, Hornell Hart's inventive contribution *Chart for Happiness* (1940) explicitly draws on both. Hart, a professor of sociology at Duke, asks: "One of the basic purposes of mankind is to be happy. Can recent advances in scientific thinking tell us more and more effectively what to do in order to be happy, and in order to help make our fellow human beings happy?" (Hart 1940, 16). He answers:

> The present book is based on two propositions: first, that it is possible to measure happiness and unhappiness reliably; and second, that, if we can thus measure, we can then move on toward discovering the causal factors by means of which we can learn with more and more effectiveness to eliminate misery and increase joy scientifically (Hart 1940, 16).

It is not clear exactly how joy and misery relate to happiness, but presumably they are relevant as important correlates of happiness. Hart notes that the development of reliable statistics about infant mortality played an important part in the discovery of the causes of infant deaths and in the subsequent reduction in infant mortality (Hart 1940, 17). Clearly, Hart hopes that something similar can be done for happiness. In this sense, the model for his research appears to be epidemiology.

Hart was also impressed by a number of measurement instruments. He opens up by discussing the uses and significance of intelligence tests, the fever thermometer and the electrocardiograph, and then goes on to say:

> *Chart for Happiness* presents and explains another such invention, the Euphorimeter, produced by the researches of various social scientists in various universities and colleges. This new instrument is to measure, not intelligence, or fever, or cardiac pulsations, but the happiness of those who submit themselves to it. It is to diagnose, not mental aptitudes, nor physical disease, but the causes of human anguish. It is to point the way, not to administration of drugs, but to constructive measures which may relieve maladjustments, promote the cure of mental suffering, and open the way toward more joyous living (Hart 1940, v).

The book is written both for those who "sense the fact that they are not living on as full a tide of happiness as they might attain" and for "psychologists, sociologists, psychiatrists, physicians, educators, social workers, pastors" and other professionals who are concerned with the happiness of others (Hart 1940, v-vi). The aim is to help people "live joyously within a menacing world ... in spite of threats and pressures of war, of economic disaster, of our own incurable physical handicaps and past emotional wounds, and of the original natures of the people with whom we have to live and work" (Hart 1940, 6).

Hart offers two different definitions of "happiness."[20] According to the first: "Happiness is the state in which people are when they say sincerely, 'I am happy,' and it is the opposite of the state in which they are when they say sincerely, 'I am unhappy'" (Hart 1940, 182). Hart adds: "It has been assumed that the average person is sincere in his answers to such a test as the Euphorimeter, especially when he does not sign his name ... Moreover, it has seemed reasonable to assume that whatever insincerities do occur tend to cancel each other out" (Hart 1940, 182). Hart notes that the first definition, under the assumption of sincerity, "serves operationally as a basis for measurement, but it is rather deficient in providing insight," and "does not take us very far into the meaning of the term" (Hart 1940, 183). Thus, he proposes a second definition: "Happiness is any state of consciousness which the person tested seeks to attain or to maintain, and it is the opposite of any state which the possessor seeks to change or from which he seeks to escape or withdraw" (Hart 1940, 183). Hart argues that the second definition is operational too (Hart 1940, 183). Hart ultimately determines that "the two definitions really define the same thing" (Hart 1940, 182-184).

The "Euphorimeter" is described in the following way:

> The scale of this instrument is somewhat like the scale of a centigrade thermometer. The zero on the Euphorimeter is the dividing point between happiness and unhappiness. From zero down, the scores mean deeper and deeper unhappiness. From zero up, the scores indicate greater and greater happiness. The 100-degree point is set by the average happiness of the general population (Hart 1940, 19).

The unit of the Euphorimeter is called a *Euphor-unit,* and is defined as "one one-hundredth of the difference between the zero point and the average happiness score" (cf. Hart 1940, 178). In fact, there are two Euphorimeters. Because people's happiness fluctuates from moment to moment, Hart writes, "it has been necessary to devise two types of Euphorimeters – one to measure happiness 'at

[20] Unfortunately, Hart does not specify whether he takes happiness to be the same thing as euphoria. The name of his measurement device certainly suggests that he identified the two.

the moment,' the other to measure as closely as possible the general level of happiness or unhappiness on which one lives in the long run" (Hart 1940, 21). Using the At-the-Moment Euphorimeter, Hart offers charts showing how the happiness of three subjects – one who was falling in love, one whose mother was dying, and one who wondered why not everybody committed suicide – evolved over time (Hart 1940, 22-24; cf. Figure 1).

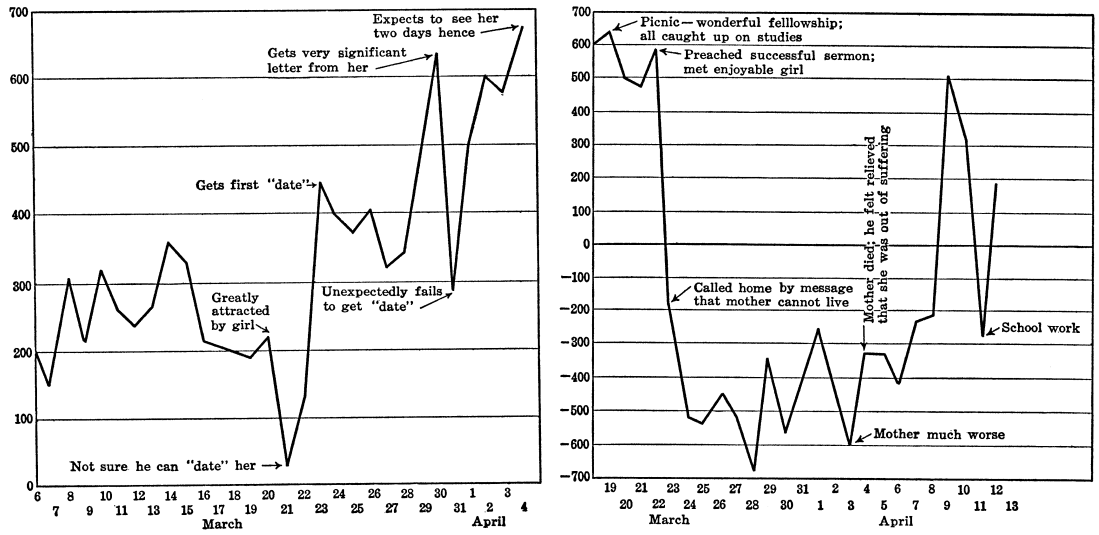

Figure 1. Happiness fluctuations in Euphor-units of a student who was falling in love (left) and of a student whose mother was dying (right) (from Hart 1940, 22-23).

The idea behind the Euphorimeter is that happiness should be measured on an "objective scale ... which will have meaning for ordinary people" (Hart 1940, 177). Apparently, Hart started out with the observation that test scores reported by his precursors all have a similar distribution.[21] In particular, Hart notes that "the scores pile up toward the happy end of the scale" and that "the distributions are all skewed toward the happy end" (Hart 1940, 175). He adds: "It seemed reasonable, therefore, to reduce the scores of all the four tests to a common scale of standard measures by subtracting from the each score the average of the group from which it comes, and

[21] Specifically, Hart discusses data presented by Hamilton (1929), Watson (1930), Sailer (1931), Terman (1938), and Burgess and Cottrell (1939) (cf. Hart 1940, 173-174).

dividing the deviation, thus obtained, by the standard deviation" (Hart 1940, 175). Furthermore, he argues, "it is necessary to have two points on the scale which are objective and which are easily understood" (Hart 1940, 176). Thus, he assigns a zero to the point that divides happiness from unhappiness, and 100 to the average happiness of all persons tested (Hart 1940, 176).

Hart's next problem is to construct an instrument that allows him to measure happiness on such a scale. For the "At-the-Moment Euphorimeter," Hart modified one of the kinds of test used by Watson (1930) and Sailer (1931). In Hart's test, the subject is offered a number of adjectives, and asked to underline each adjective "which fairly well describes the way you have felt a good deal of the time or several times since you woke up this morning" as well as to cross out each adjective "which does NOT describe the way you have felt at any time today" (Hart 1940, 114). The adjectives were chosen by picking all synonyms and antonyms of "happy" and "unhappy" in a dictionary, then adding all synonyms and antonyms of those, and so on, until he had a list of 48 adjectives (Hart 1940, 176-177). A subject is said to have a happy reaction each time he or she underlines a positive adjective or crosses out a negative one, and an unhappy reaction each time he or she underlies a negative adjective or crosses out a positive one (Hart 1940, 115-116). The test score can then be computed: "The adjective quotient is derived by subtracting the number of unhappy reactions from the number of happy reactions and dividing this difference by the sum of happy plus unhappy reactions" (Hart 1940, 177). Thus, if the subject has an equal number of happy and unhappy reactions, the score will be zero. After administering the test to 2,200 subjects, Hart found that the average quotient was approximately .25 (Hart 1940, 178). Thus, the adjective quotient has to be multiplied by 400 in order for the average happiness score to equal 100. Hart assures himself that the test is reliable by computing scores separately for each half of the adjectives, and finding the reliability of the entire test to be .90 (Hart 1940, 179).

Similarly, for the "Long-Run Euphorimeter" Hart borrowed a number of questions from earlier studies and added some "which seemed likely to be valuable" (Hart 1940, 181). For many of the questions on this test, the subject is asked to consider how they "usually" feel, rather than how they feel at the moment (Hart 1940, 111). The scoring system for the Long-Run Euphorimeter was worked out by comparing the results of 2,200 subjects who took both the long-run and the at-the-moment test. More specifically, Hart says that "scoring values were worked out in such a way that any large group of persons who each score approximately zero on the adjective battery will average approximately zero on the Long-Run Euphorimeter," and so on (Hart 1940, 181).

The validity of the measures is assured, Hart maintains, by the manner in which he has defined "happiness." In effect, he argues, "the adjective battery consists in offering an opportunity to say 'I am happy' in 48 different ways, or to say 'I am unhappy' in 48 different ways, or to make whatever compromise between these extremes fit one's own state at the time tested" (Hart 1940, 182). The Long-Run test also contains questions concerning how subjects feel about the change or maintenance of current conditions. Hart claims:

> At every crucial point in a long series of statistical analyses, based on these data, we have found that persons that say, in one form or another, "I am happy," tend also to say, in various ways, "I want to keep my present way of life developing as it is now going" ... The correlation is so close that it seems wisest to proceed upon the assumption that our two definitions of happiness really define the same thing (Hart 1940, 184).

He concludes:

> Various improvements need to be made, and will continue to be made, in the Euphorimeter tests. But it is believed that, as they now stand, with their present scoring methods, the Euphorimeters measure happiness and unhappiness with sufficient reliability and validity to identify outstandingly happy and unhappy people and to give important aid in the process of increasing the happiness of those tested (Hart 1940, 184-185).

In brief, Hart offers two new definitions of happiness, and new tools to measure it. The fact that he offers two different definitions appears to reflect a certain vacillation on his part. The definitions offered have more of an operational flavor than the others considered so far. As far as I can tell, however, he never makes up his mind on which definition is the right one; in Hart's view, the claim that the two "define the same thing" – by which he means, I suppose, that they have the same extension – obviates the need to make a choice. When it comes to Hart's methods to measure happiness, he relies on a modified version of tests that have been used by others. His main conceptual innovation is the Euphorimeter, and the scale on which to make measurements.

In 1954 Herbert Jeremy Goldings of the Harvard Psychological Clinic offered another study which, following Hart (1940), drew on research both on marital happiness and educational psychology. Goldings' article is interesting in part because it uses new methods to measure happiness, and it contains a more explicit discussion about the notion of happiness that the author has in mind. Like many others in this tradition, Goldings remarks on the lack of attention to questions of happiness on the part of professional psychologists. He writes that "there has been a marked reluctance on the part of present-day researchers to undertake systematic studies of happiness, unhappiness, and related phenomena" (Goldings 1954, 30). He adds: "By contrast, there is a commercially vigorous popular concern with happiness (books, magazines, etc.) indicating, perhaps, that in American culture people are not happy, or not as happy as they want to be, or not as happy as they feel they should be" (Goldings 1954, 30). Goldings remark indicates that commercial culture has not changed much during the last half-century.

Goldings starts out by considering "the intrinsic nature and characteristics" of happiness (Goldings 1954, 31). He writes: "Happiness and unhappiness may be considered as zones on a continuum of hedonic affect which embraces feelings of elation, contentment, satisfaction, and

pleasure at the positive pole and feelings of depression, discontent, and unpleasure at the negative pole" (Goldings 1954, 31). He adds that

> ... although, in common parlance, "happiness" refers only to the positive elements of hedonic feeling, in the present investigations "happiness" more generally refers to the entire continuum as a whole and "satisfaction" and "dissatisfaction" refer more specifically to the polar regions of positive and negative hedonic feeling (Goldings 1954, 31).

From this quote it appears that Goldings draws a distinction between "happiness" – the continuum of hedonic feeling – and "satisfaction" – the positive end of this continuum. It is unclear whether he in fact uses the terms in this way, however. Consider:

> In order to bring happiness and unhappiness (or satisfaction and dissatisfaction) into proper perspective *vis-a-vis* other general areas of psychological study, they may be considered hypothetical physiological states with both *subjective manifestations* (diffuse and pervasive feelings of pleasantness and satisfaction, and of unpleasantness and dissatisfaction) and *objective signs* (posture, gait, facial expression, tone of voice; feelingful verbal statements; indirect or projective indices of feeling) (Goldings 1954, 31, italics in original).

Here, happiness and satisfaction are treated as physiological states that cause both subjective feelings, and outward behaviors, of a certain kind.

Goldings had 20 subjects, all of whom were Harvard undergraduates. First, subjects were shown 30 photographs of faces with ambiguous facial expressions, and asked "to rate the satisfaction or dissatisfaction (happiness or unhappiness) of the people" (Goldings 1954, 34). He added: "The *Ss* [subjects] rated each picture twice on a 10 point scale (extremely unhappy to extremely happy). The sum of the 60 ratings (2 ratings for each of the 30 pictures) constituted the *S's ascribed-happiness* score" (Goldings 1954, 34, italics in original). It remains unclear if Goldings asked once about a person's happiness and once about his or her satisfaction, or if he asked the same question (phrased as above) twice. The reason why Goldings included this test is that he feared "direct avowal of happiness may tend to be subjected to some distortion," and that it would be

important to have alternative, indirect, means of measuring happiness. The significance of the happiness ascribed to others with ambiguous facial expressions is that the subject would be expected, on theoretical grounds, to project his or her own happiness onto the subject (Goldings 1954, 33).

Second, subjects were given an adaptation of the test used by Watson (1930) and Sailer (1931): "The form ... consisted of linear-type rating scales for the avowal of the S's happiness or unhappiness in eight areas of life (...), and a ninth scale for general, over-all satisfaction" (Goldings 1954, 34-35). The subject were asked to "think of his feelings over the last year" when estimating his own happiness, so as to obtain "what is, perhaps, an estimate derived from a series of changing hedonic states over a period of time at least more extensive than the immediate present" (Goldings 1954, 32). On the basis of the results, Goldings computed an *avowed happiness* score ranging from 1 to 6 (Goldings 1954, 35). Third, Goldings asked five experimenters to rank-order the subjects "on 'general happiness and overall satisfaction' based on the experimenter's subjective, clinical judgment of the individual" (Goldings 1954, 35). The ratings were retrospective, and based on the impression made by the subjects during previous testing periods (Goldings 1954, 35).

Among other things, Goldings confirmed Watson's finding that subjects tend to rate their own happiness as greater than the average (Goldings 1954, 36). Goldings thus reaffirmed his hypothesis that there is a general "tendency to avow or overavow happiness and disavow or underavow unhappiness" (Goldings 1954, 46). Moreover, he found that there was "highly significant agreement among the five judges," and that there was "a fairly close general agreement between the *S's* avowal of his own happiness and the ratings assigned by the clinical judges" (Goldings 1954, 40).[22] In contrast, he found no positive correlation between avowed happiness and projected

[22] Goldings also notes that there was more agreement about the unhappy subjects than about the happy ones (Goldings 1954, 41).

happiness. A closer examination of the data suggests to Goldings that those subjects whose avowed happiness were in the normal range indeed tended to attribute their own mental state to the ambiguous faces, whereas those subjects who were extremely happy or unhappy tended to attribute the opposite state (Goldings 1954, 42).

In summary, Goldings is more explicit about what he means by "happiness" and "satisfaction" than many other authors. According to his definition, both refer to physiological states with generate both characteristic mental states and outward behaviors. Unfortunately, there appears to be a disconnect between the concepts that he defines and the concepts that he actually uses in his research. Thus, it remains unclear how happiness and satisfaction are defined. Goldings experiments with three different methods to measure happiness, and finds that avowed happiness correlates with the judgments of experienced judges, but not with projected happiness. Although he does not explicitly spell out the implication, he clearly concludes that avowed happiness and the ratings of experienced judges, but not projected happiness, can serve as measures of happiness.

2.2.5 The epidemiology of mental health

In the late 1950's and early 60's, a number of mental health professionals began using measures of happiness and satisfaction on a large scale. A landmark in the area was *Americans View Their Mental Health*, by Gerald Gurin, Joseph Veroff, and Sheila Feld (1960). As the title indicates, Gurin et al. were interested in how Americans themselves – as opposed to mental health experts – see their mental health, but also in how likely they are to seek professional help (Gurin et al. 1960, 3-4). Gurin et al. adopt a multiple criterion perspective, that is, they believe that mental health is best measured by using a battery of criteria. One such criterion is whether people feel happy. Unfortunately, Gurin et al. say little about what they mean by the term, beyond discussing what subjects think contributes to happiness (Gurin et al. 1960, 22-28). One of the questions they asked their sample of 2,460

randomly selected non-institutionalized American adults (Gurin et al. 1960, 3) was the following: "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, 411, italics in original).

On the basis of their research, Gurin et al. found that 35 percent of respondents described themselves as "very happy," 54 percent as "pretty happy," and 11 percent as "not too happy" (Gurin et al. 1960, 22). They found an association between unhappiness and pessimism and uncertainty about the future (Gurin et al. 1960, 35), but no connection between the sex of the subject and his or her happiness (Gurin 1960, 42). They found that older people were less happy than younger (Gurin et al. 1960, 44), and that those with more education were happier than those with less (Gurin et al. 1960, 46). They conclude: "Education, like youth, seems to be associated with the investment of greater aspirations and expectations in life – an investment which brings greater gratification" (Gurin et al. 1960, 51). They also find that happiness is positively related to income (Gurin et al. 1960, 216). Thus, Gurin et al. (1960) do not identify happiness and mental health, but see the former as an important component of the latter. They use a direct question to assess the degree of happiness of their subjects. As we will see later on, the question has been reused in a great number of subsequent studies.

Norman M. Bradburn and David Caplovitz's book *Reports on Happiness: A pilot study of behavior related to mental health* (1965) built on Gurin et al. (1960) and is another important study in the field.[23] According to Bradburn and Caplovitz, their research "is an effort to develop, for psychological and behavioral phenomena, time-series studies comparable to those that are commonplace in economics and demography" (Bradburn and Caplovitz 1965, 1). They continue:

> Its long-range objective is to conduct periodic inventories of the psychological well-being of the nation's population. From such inventories it is possible to determine for the first time the extent to which feeling states of

[23] Bradburn (1969) report more results from the same study.

> the population are affected by major social trends, national and local crises, and changes in the economic and social structures, as well as by patterned events in the life cycles of individuals (Bradburn and Caplovitz 1965, 1).

The assumptions underlying their research project are relatively explicit. Bradburn and Caplovitz write:

> The underlying assumption of this research is that there is a dimension variously called mental health, subjective adjustment, happiness, or psychological well-being, and that individuals can be meaningfully described as being relatively high or low on such a dimension (Bradburn and Caplovitz 1965, 1)

Although the passage is tentative, it does suggest that the authors identify mental health, happiness and psychological well-being, and they clearly take it to be a uni-dimensional entity. In their view: "The most fruitful starting point in developing an instrument to measure a dimension of mental health seemed to be people's own estimates of their level of psychological well-being or distress" (Bradburn and Caplovitz 1965, 5). According to Bradburn and Caplovitz, the degree of psychological well-being may as well be assessed by simple self-report measures. They write that

> ... there is no evidence that self-reports are any less (or for that matter more) valid than expert ratings or psychological tests for rating people on a mental health dimension. Furthermore, self-reports have the eminently practical virtues of face validity, directness, and ease of use (Bradburn and Caplovitz 1965, 7).

The authors are aware, however, that there is no agreement on what to call the dimension of interest, or on how it should be measured (Bradburn and Caplovitz 1965, 1).

Bradburn and Caplovitz interviewed or administered questionnaires to 2,006 members of 450 households in four rural Illinois communities (Bradburn and Caplovitz 1965, 3-5). In order to measure the degree of happiness of the respondents, Bradburn and Caplovitz used the question articulated by Gurin et al. (1960), viz. "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?" (Bradburn and Caplovitz 1965, 146). The authors' results – 24 percent said they were "very happy,"

59 percent "pretty happy," and 17 percent "not too happy", respectively – were slightly different from those of Gurin et al. (1960), which is unsurprising given the different sample (Bradburn 1965, 8). Other than that, the authors report that their findings were similar to those of Gurin et al. (1960) (see Bradburn and Caplovitz 1965, 10). In summary: "happiness is positively correlated with education and income, negatively correlated with age, and uncorrelated with sex (Bradburn and Caplovitz 1965, 56).

Perhaps the most surprising conclusion of Bradburn and Caplovitz is that positive affect is unrelated to negative affect. In order to develop "more detailed measures of well-being," Bradburn and Caplovitz offered twelve different descriptions – half positive, half negative – and asked subjects how often they had felt that way during the past week (Bradburn and Caplovitz 1965, 15-16). As they write: "We expected that the items would cluster in two groups, one indicative of positive and the other of negative feelings, and that the two clusters would be negatively related to one another" (Bradburn and Caplovitz 1965, 16). As expected, the authors did find "a strong tendency for most of the items to fall into two clusters of positive and negative feelings" (Bradburn and Caplovitz 1965, 16). However, they were surprised to learn that "items in one cluster are not negatively related to those in the other cluster in any consistent or strong fashion" (Bradburn and Caplovitz 1965, 18). The two items which were the most strongly (negatively) correlated were "depressed or very unhappy" and "on top of the world," and these items had a correlation of a mere -.19 (Bradburn and Caplovitz 1965, 18). Interestingly, while positive and negative items do not correlate with each other, each is correlated with happiness (Bradburn and Caplovitz 1965, 18-19). The authors infer that "happiness is a result of the relative strengths of positive and negative feelings, rather than of the absolute amount of one or the other" (Bradburn and Caplovitz 1965, 21). In other words, the

presence of strong negative feelings during some period of time need not compromise one's happiness, if there are at the same time sufficiently strong positive feelings.[24]

2.2.6 Review and call for more research

In 1967, Warner Wilson offers the first review article of research on happiness, or as he says, borrowing Goldings' term, "avowed happiness." The article is based on his doctoral dissertation. Wilson reviews the research discussed above, his own dissertation work, and more. On the basis of his meta-analysis, Wilson concludes that "avowed happiness can be determined reliably," and that the "facts ... support the validity of self-ratings" (Wilson 1967, 294-295). On the basis of the studies he reviews, Wilson infers: "The happy person emerges as a young, healthy, well-educated, well-paid, extroverted, optimistic, worry-free, religious, married person with high self-esteem, high job morale, modest aspirations, of either sex and of a wide range of intelligence" (Wilson 1967, 294).

Wilson ends his article by calling for more experimental research, in which the happiness of subjects is manipulated through therapy or other means. He writes:

> A story is told of how a small group of college men increased the poise and popularity, and presumably the happiness, of a female student by going out of their way to respond to her as though she were attractive. The possibility of systematically developing and applying such principles and techniques seems exciting indeed! (Wilson 1967, 305).[25]

This is the last time (to my knowledge) that somebody seriously suggested this kind of experiment. Perhaps ethical constraints on scientific research put an end to it.

---

[24] In this context, I might mention Alden E. Wessman and David E. Ricks (1966), who examine the degree of happiness and unhappiness in the context of a study of mood and personality.

[25] References have been omitted.

2.2.7 Gerontology

Meanwhile, a number of studies on subjective well-being had appeared in the field of gerontology.[26] According to a David L. Adams, a concern with "individual well-being" was one of the factors contributing to the development of the field in the first place (Adams 1971, 64). He continues: "However, the difficulty of trying to assess 'individual well-being' has resulted in a variety of concepts, definitions, and measurements" (Adams 1971, 64). Some of the terms that have been used to assess well-being in gerontology include "morale," "successful aging" and "personal adjustment." According to Bernice L. Neugarten, Robert J. Havighurst, and Sheldon S. Tobin (1961), from the start there were two different approaches to measuring well-being in gerontology. One examined overt behavior, and compared it with what was taken to be socially accepted criteria of "success" or "competence" (Neugarten et al. 1961, 134). This approach was criticized, however, for relying too heavily on the assumptions that social competence (and participation) is closely associated with well-being, and for imposing the value judgments of the investigators on the subjects (Neugarten et al. 1961, 134).

At least in part in reaction to such criticism, some investigators adopted another approach. In this approach, the focus is on the "internal frame of reference," in which qualities like social competence figure only indirectly (Neugarten et al. 1961, 134). As they write:

> Here the variables to be measured have been the individual's own evaluations of his present or past life, his satisfaction, or his happiness. The assumptions are, whether or not explicitly stated, that the individual himself is the only proper judge of his well-being; that the value judgments of the investigator can thus be minimized; and, perhaps most important, that it is not appropriate to measure well-being in old age by the same standards that apply to middle age, namely, standards based upon activity or social involvement (Neugarten et al. 1961, 134).

[26] For reviews of this literature, see especially Neugarten et al. (1961), Adams (1971), and Larson (1978).

In this view, gerontologists took the subjective turn in order to avoid imposing their own standards of well-being and, their own value judgments, on their subjects. Apparently, they wanted to rely (insofar as possible) on the subjects' own standards of what constitutes a life of well-being. Moreover, they felt that the subjects themselves were best situated to determine to what extent they satisfied those standards.

There were several studies that relied on the "internal frame of reference" (e.g. Kuhlen 1948; Lebo 1953). Many of the early studies combined elements of the two approaches, and asked questions about "happiness" and "feelings of well-being or satisfaction" along with questions designed to assess degrees of social competence (Pollak 1948, 66-67). Thus, in an influential 1949 book, Ruth Shonle Cavan, Ernest W. Burgess, Robert J. Havighurst, and Herbert Goldhamer (1949) aspire "to define and to analyse the nature, patterns, and problems of personal adjustment to ageing" (Cavan et al. 1949, v).[27] They asked their 3000 participants questions both about whether they feel satisfied, and whether they participate in clubs and organizations (Cavan et al. 1949, Appendix A). This study, like that of Pollak (1948), came out of a research project on "the social aspects of old age" organized by the Social Science Research Council (Cavan et al. 1949, vi; see also Havighurst 1957).

The most common focus in gerontology, however, is on "life satisfaction" (cf. Neugarten et al. 1961; Adams 1971). Thus, Arnold M. Rose (1955), who appears to have been a sociologist at the University of Minnesota, studied life satisfaction of married adults whose children had just left the home. Rose asked his subjects: "In general, how satisfied are you with your life?" and offered them a list of five alternative answers: "Very dissatisfied," "Somewhat dissatisfied," "About average," and

[27] One of the authors of this book, Ernest W. Burgess, was also the co-author of Burgess and Cottrell (1939), which was a major contribution to the literature on marital happiness. This, the literature developed in this section was not entirely independent from that discussed earlier.

so on (Rose 1955, 15).[28] Rose found that women (but not men) who married before the age of 20 or after the age of 30 tend to be less satisfied with their lives than others, that dissatisfied women are less likely to have a paying job and more likely to think they spend an inordinate amount of time doing housework, and that satisfaction is associated with social participation in both men and women (Rose 1955, 18-19).

A more explicit treatment of measures of life satisfaction appears in Neugarten et al. (1961), all of whom were affiliated with the Committee on Human Development at the University of Chicago. Unfortunately, these authors do not say much about what they mean by satisfaction. They indicate that they take satisfaction to be closely related to "Zest (vs. apathy); Resolution and fortitude; Congruence between desired and achieved goals; Positive self concept; and Mood tone" (Neugarten et al. 1961, 137), but do not specify if these factors are supposed to be components, causes, or mere correlates of satisfaction. Anyway, Neugarten et al.'s main contribution is the method they propose to measure life satisfaction. In fact, they constructed two indices of life satisfaction. The first (*Life Satisfaction Index A*) consisted of 20 statements of the form "I've gotten pretty much what I expected out of life" and checkboxes for "AGREE," "DISAGREE" and "UNSURE" (Neugarten 1961, 141). The score was computed by awarding the respondent one point every time he or she agreed with a statement indicating satisfaction or disagreed with one indicating dissatisfaction. The second index (*Life Satisfaction Index B*) consisted of twelve questions such as "How satisfied would you say you are with your way of life?" and three options "very satisfied" (for two points), "fairly satisfied" (one point) and "not very satisfied" (zero points). The score was computed by adding up the points in parentheses. The two indices were supposedly validated by comparing them to the score (*Life Satisfaction Rating*, or LSR) assigned to the respondents by

[28] It is unclear if, or to what extent, Rose (1955) was inspired by the marital happiness literature surveyed in the previous chapter.

experienced judges in structured, lengthy interviews. The investigators found that correlation coefficient between LSIA correlated .55 with LSR, and LSIB correlated .58.

2.2.8 The social indicators movement

A great deal of interest in subjective well-being during the 1960's and 70's appears to stem from the social indicators movement.[29] This movement emerged in the late 1960's as a reaction against the widespread use of economic indicators like GDP as measures of "the goodness of life," (Campbell 1976, 118; Andrews 1989, 401). Campbell admits that economic measures are "easy to count," and he suggests that this goes a long way toward accounting for their popularity as measures of well-being. Yet, he adds: "None of us doubts that economic data have admirable qualities; the question is, How well do they represent the quality of national life? How valid are they as measures of the goodness of life in this country?" (Campbell 1976, 117). Campbell himself, obviously, is not enthusiastic about economic measures.

In contrast, the social indicator movement attempted to offer a set of indicators such that, although perhaps not as "easy to count," would nevertheless be better measures of the quality of life. As Andrews (1989) puts it:

> The idea called for two key changes from earlier practices. One was an expansion in the range of phenomena monitored beyond the traditional economic indicators, and an explicit recognition that "life quality," however it might be designed, involved more than just economic considerations. The second change involved an attempt to focus directly on "output" indicators – i.e., indicators that show how well off people actually are – in addition to the more traditional "input" indicators that reflect budget allocations, procedures and processes that are presumed to enhance well-being (Andrews 1989, 401).

[29] See section 1.2.3 for slightly more information of the origins of this movement.

Thus, members of this movement encouraged the collection of data on life expectancy, quality of food and water, access to adequate medical care, level of education, quality of housing, and so on. Such indicators are often referred to as "objective indicators" (Campbell 1976, 118). As Campbell notes: "It is reasonably argued that as the level of education rises, the adequacy of medical care improves, the amount of substandard housing is reduced, and the purity of the air and water is increased, the quality of life is therewith enhanced (Campbell 1976, 118)." The underlying assumption, of course, is that more education, housing, and medical care tend to increase one's quality of life.

Nevertheless, it appears that a number of researchers felt that developing indicators of people's level of education, quality of housing, and access to medical care did not go far enough. As Campbell puts it:

> [If] we believe, as I assume that most psychologists do, that the quality of life lies in the experience of life, then these are surrogate indicators. They describe the conditions of life that might be assumed to influence life experience, but they do not assess that experience directly (Campbell 1976, 118).

In Andrews' terms, indicators that concern education, housing and health care remained "input" indicators. Thus, the very argument that led to the rejection of economic indicators appears to have led to the rejection of "objective" indicators as well. In Campbell's words:

> If we are primarily concerned with describing the quality of life experience of the population, we will need measures different from those that are used to describe the objective circumstances in which people live. We will have to develop measures that go directly to the experience itself. These *subjective* measures will surely not have the precision of indicators that are expressed in numbers of dollars, units of time, or numbers of square feet, but they will have the great advantage of dealing directly with what it is we want to know, the individual's sense of well-being (Campbell 1976, 118, italics in original).

It is clear that the social indicator movement was motivated primarily by a concern with the quality of life or sense of well-being – the terms are not clearly distinguished from each other – from the

very beginning. Moreover, some clearly felt that the logic of their position forced them to reject the exclusive reliance not only on economic measures, but on a wider set of objective measures as well.

One important study is Princeton psychologist Hadley Cantril's book *The Pattern of Human Concerns* (1965). Cantril is explicit that his research is aimed not only to advance scientific understanding, but also to design "new institutions, new political, economic, and social mechanisms" (Cantril 1965, 3). He adds: "More reliable predictions of what people want or do not want, believe or do not believe, will accept of will not accept, should aid the process of creating new forms of economic, social, and political institutions" (Cantril 1965, 3). To make such predictions, Cantril is interested in people's aspirations, and the degree to which those aspirations are satisfied. His goal is to find a method to estimate the aspirations and satisfactions by reference to the subjects' own, possibly changing standards rather than his own (Cantril 1965, 21).

Thus, Cantril invented the "Self-Anchoring Striving Scale" (Cantril 1965, 22). When using the scale, the subject is asked to consider the best and worst possible life he or she could live, and to judge how good his or her life is by comparison to the best and worst possible life, thus described. Specifically, the subject is asked: "When you think about what really matters in your own life, what are your wishes and hopes for the future? In other words, if you imagine your future in the *best* possible light, what would your life look like then, if you are to be happy?," and similarly for the *worst* possible life (Cantril 1965, 23). Thereafter, the subject is shown a non-verbal advice referred to as "the ladder of life," a drawing of a ladder-like shape where the rungs have been numbered from 0 to 10 (Cantril 1965, 22). Next, the subject is told: "Here is a picture of a ladder. Suppose we say that the top of the ladder (POINTING) represents the best possible life for you and the bottom (POINTING) represents the worst possible life for you," and asked "Where on the ladder (MOVING FINGER RAPIDLY UP AND DOWN THE LADDER) do you feel you personally stand at the *present* time?"

(Cantril 1965, 23).[30] The subject is asked to answer by giving the number of the step (Cantril 1965, 23). Cantril's basic idea is that this device will allow him to judge a person's circumstances from the person's own point of view. In this test, Cantril writes: "A person is asked to define on the basis of *his own* assumptions, perceptions, goals, and values the two extremes or anchoring points of the spectrum on which some scale measurement is desired" (Cantril 1965, 22). To his surprise, Cantril learned that in some cultures ladders are uncommon. Thus, in some places he substituted a picture of a mountain for the ladder (Andrews and Robinson 1991, 73).

Under the heading "Who Are the Satisfied?," Cantril discusses what characterizes individuals who rate themselves high on the Self-Anchoring Striving Scale. Cantril and his collaborators administered the test to a total of 23,875 subjects in 13 countries. Among other things, Cantril finds that "education, income, and occupation" are strongly correlated with ladder ratings, that "people living in urban centers" rate themselves somewhat higher than "those living on the land," but that "men and women rate themselves about the same" (Cantril 1965, 258).

In certain ways, Cantril offers a different approach than the other authors examined so far. Of course, his use of a graphical rating device to elicit responses is not novel, though Cantril relies on a vertical ladder (or a mountain) rather than a horizontal bar like Watson (1930). Cantril's study is novel, however, first in that it explicitly focuses on satisfaction rather than happiness, and because it uses the author's self-anchoring technique for the subjects. This is an interesting device to help respondents assign meaning to the numbers they have to provide the experimenter, but introduces a number of problems of interpretation (as Cantril notes; cf. 1965, 25).

Angus Campbell, Philip E. Converse, and Willard L. Rodgers open up their discussion in *The Quality of American Life* (1976) by pointing to deficiencies both with traditional economic measures of well-being, and with those proposed by the social indicator movement. The authors suggest that

[30] The parenthetical notes in small caps are Cantril's instructions to the experimenter (Cantril 1965, 22).

there may be a trend away from such measures when they note that "those who presume to define the national goals increasingly speak of quality of life rather than of further material possessions" (Campbell et al. 1976, 1). They add:

> Quality of life is seldom precisely defined in these statements, but the implication is typically given that the nation must change from its fixation on goals which are basically economic to goals which are essentially psychological, from a concentration on being well-off to a concern with a sense of well-being (Campbell et al. 1976, 1).

The social indicator movement, Campbell et al. claim, is concerned with developing measures of the quality of life other than the established economic ones (Campbell et al. 1976, 2-3). Thus, "measures which ... deal with population growth and movement, marital status, unemployment and labor-force participation, health and health care, housing, education, leisure time, crime, and the administration of justice," according to Campbell et al. "are commonly taken as measures of the achievements and well-being of a society" (Campbell et al. 1976, 3). In the process, they note, "these measures of the objective conditions of life are taken as surrogates for the subjective experience of life" (Campbell et al. 1976, 3).

Although the translations from objective conditions to subjective experience may appear plausible, "the fact is that we do not know how well objective measures like these represent underlying psychological states or how well social indicators can be taken to represent the quality of life experience" (Campbell et al. 1976, 3). The 1976 book was written to make up for this deficiency:

> The research with which this book is concerned derives from the conviction that the relationship between objective conditions and psychological states is very imperfect and that in order to know the quality of life experience it will be necessary to go directly to the individual himself for his description of how life feels to him. This obviously will take us into the subjective world of perceptions, expectations, feelings, and values and will involve us in excruciating problems of definition and measurement (Campbell et al. 1976, 4).

The book has two objectives, Campbell et al. say. First, the authors begin to establish "a system of reporting of psychological data which parallels our current repertoire of social indicators" (Campbell et al. 1976, 4). Second, Campbell et al. want to provide "a fuller and truer representation of the state of society to those people who are responsible for social decisions" (Campbell 1976, 5).

Campbell et al. say they drew primarily on the three following sources: Gurin et al. (1960), which they refer to as the "first major study of the quality of life experience," Bradburn and Caplovitz (1965), and Cantril (1965) (see Campbell et al. 1976, 5). In their own study, Campbell et al. decided to use a measure of satisfaction – as opposed to happiness – as a measure of the quality of life experience (Campbell et al. 1976, 7-8). First, they took "satisfaction" to be easier to define and to translate than "happiness" (Campbell et al. 1976, 7-8). Second, the authors wanted their data to be relevant to public policy, and they thought legislators and decision-makers were more accustomed to thinking in terms of satisfying needs than to promote happiness (Campbell et al. 1976, 8-9). Third, Campbell et al. wanted a concept appropriate to use both in the context of "life as a whole" as well as in the context of more limited domains (Campbell et al. 1976, 9; cf. p. 33).

Campbell et al. offer an explicit definition of the term "satisfaction." They write:

> Level of satisfaction can be precisely defined as the perceived discrepancy between aspiration and achievement, ranging from the perception of fulfillment to that of deprivation. Satisfaction implies a judgmental or cognitive experience, while happiness suggests an experience of feeling or affect (Campbell et al. 1976, 8).[31]

In this context, Campbell et al. point out that reports of happiness and satisfaction do not correlate perfectly. Typically, they say, the correlation is about .50 (Campbell et al. 1976, 8). They write: "This means that there is some significant minority of persons who report relative happiness along with

[31] There seems to me to be something backward about this definition. If the perceived discrepancy between aspiration and achievement goes up, presumably satisfaction should go down, but the definition suggests the opposite.

relative dissatisfaction with their lives, and the converse" (Campbell 1976, 9). The lack of correlation also means that the choice of measure may affect the conclusions of a study.

Campbell et al. note that the degree of satisfaction may differ radically across individuals even under objectively identical circumstances. They also note that this fact is part of the *raison d'être* of the whole endeavor: "If there were a close and universal relationship between the level of material possessions and the quality of life experience, there would, of course, be little point in undertaking a study of the kind in which we are here engaged" (Campbell et al. 1976, 10). Meanwhile, the authors offer an interesting discussion of the various factors that may affect a person's degree of satisfaction. They write that satisfaction depends on two things, viz. "how he perceives the attribute and the standard against which he judges that attribute" (Campbell et al. 1976, 14). Note the emphasis on the individual's perception of attributes, as opposed to the attributes themselves. If people's perceptions were always accurate, this distinction would make little difference. However, if people's perceptions sometimes are inaccurate, as appears likely, the distinction matters. Campbell et al. (1976) continue:

> The concept of a standard of comparison or a frame of reference for such judgments is admittedly difficult to define and probably depends on multiple criteria at once. The individual's assessment may derive from any or all of the following bases of evaluation: *aspiration levels*, or the situation that a person hopes eventually to attain, where a given domain is concerned; *expectation levels*, or the situation he feels he is likely to attain in the fairly immediate future; *equity levels*, or what he thinks should be true of his situation if perfect justice prevails, given how much he invests in it relative to others; *reference group levels*, or what he believes to be true of the situations of others with whom he identifies, such as friends and family or others of his income, race, or occupation; *personal needs*, or the amount of a particular reward he may require, such as how much savings to feel secure, how much housing to be comfortable, how much police protection to feel safe; and *personal values*, concerning such intangibles as freedom, equality, and the like (Campbell et al. 1976, 14).[32]

[32] There are interesting references to Maslow's hierarchy of needs in some of these passages.

This is a nice characterization of the various factors – leaving out perceived achievements – that affect satisfaction. Campbell et al. also speculate that satisfaction may be affected by such comparisons to a greater degree than happiness, though surely both are to some degree (Campbell et al. 1976, 31).

As indicated above, the authors of this study were interested in asking questions about specific domains, concerning subjects' satisfaction with their marriage, job, and housing, etc. However, the authors also asked questions about subjects' satisfaction with "life in general," and their happiness (using the question in Gurin et al. 1960) (Campbell et al 1976, 13). Finally, the authors gave subjects "a series of descriptive adjectives" by which they were supposed to describe their lives, and used the results to compute what they called an "Index of General Affect" (Campbell et al. 1976, 13). The idea was to explore how domain satisfaction relates to general life satisfaction, and how the latter relates to other measures of the "quality of life experience." For a great many of the comparisons, however, Campbell et al. rely on a composite "Index of Well-Being," which is "a single variable representing a rather global sense of well-being," and which combines the satisfaction score and the result on the adjective test Campbell et al. estimate that the reliability of the Index of Well-Being exceeds .53 "by some significant amount" (Campbell et al. 1976, 49).

On the basis of personal interviews with a random sample of 2,164 Americans living in households (Campbell et al. 1976, 511), the authors report: "Persons who are currently single generally report a good deal less satisfaction with life than do married persons, and the lack of satisfaction shown by women and men who are divorced or separated is quite remarkable indeed" (Campbell et al. 1976, 36). Moreover, the young tend to report higher happiness than the old, but interestingly the old appear to be more satisfied with life than the young (Campbell et al. 1976, 36). The unemployed, like the divorced, score "conspicuously low," and there is a "rather strong and regular relationship between income and a sense of well-being" (Campbell et al. 1976, 51-55).

Incidentally, these relationships are discussed in part to "give a first sense of" the validity of the measure (Campbell et al. 1976, 57).[33]

Frank M. Andrews and Stephen B. Withey's book *Social Indicators of Well-Being: Americans' perceptions of life quality* (1976) is another major contribution. Andrews and Withey place their research firmly in the social indicator movement. They write that "social indicators of the various conditions of human beings, and of the changes that characterize their lives, offer much that is attractive," but add that "in these days of growing interdependence and social complexity we need more adequate cues and indicators of the nature, meaning, pace, and course of social change" (Andrews and Withey 1976, 1). Moreover, they add that "most social phenomena are of our own making and subject to our own direction" (Andrews and Withey 1976, 2). Of everything that could be studied under this heading, Andrews and Withey choose to concentrate on "perceptions of well-being" (Andrews and Withey 1976, 6; cf. 10). They write:

> Measurements of individual well-being seem to us, and to many others also, a particularly promising place to begin. The promotion of individual well-being is a central goal of virtually all modern societies, and of many units within them. While there are real and important differences of opinion – both within societies and between them – about how individual well-being is to be maximized, there is nearly universal agreement that the goal itself is a worthy one and is to be actively pursued (Andrews and Withey 1976, 6-7).[34]

According to these authors, the importance of measuring well-being derives from the importance of promoting it in a well-informed manner.

In their study, Andrews and Withey test 68 different measures, but settle on one that they think has more desirable qualities than the others (Andrews and Withey 1976, chapter 6). This is the *Delighted-Terrible Scale* (D-T Scale), which consists of seven categories ranging from "Delighted,"

[33] In the very last chapter, Campbell et al. offer an interesting discussion about the limitations of policy relevance.

[34] A reference has been omitted.

though "Pleased," and "Mostly satisfied," to "Mixed (about equally satisfied and dissatisfied)" to "Mostly dissatisfied," "Unhappy" and "Terrible" (Andrews and Withey 1976, 18). The subjects are given three additional options, viz. "Neutral (neither satisfied nor dissatisfied)," "I never thought about it," and "Does not apply to me" (Andrews and Withey 1976, 18-19). Subjects were told to consider "the feelings you have now," but were also encouraged to take "into account what has happened in the last year and what you expect in the near future" (Andrews and Withey 1976, 19).

Andrews and Withey (1976) address head on the question of whether it is reasonable to assume that people make the kind of global judgments that their research presupposes. One of the reasons why they think it is meaningful to ask for global assessments is that people articulate such judgments "promptly and with apparent ease" (Andrews and Withey 1976, 64). Moreover, in their sample, less than one percent checked the box marked "Never though about it" when asked how they feel about their life as a whole (Andrews and Withey 1976, 64). Finally, they suggest, the fact that questions such as "How are things" are so common suggest that people think in those terms (Andrews and Withey 1976, 64-66).[35]

In sum, there is a good amount of evidence that research on subjective well-being was influenced by the social indicator movement. It is interesting that two of the main proponents of the social indicator movement, Frank M. Andrews and Angus Campbell, were simultaneously two of the most important contributors to the literature on subjective well-being. Andrews and Campbell's contributions, by the way, appeared during the 70's, which can be characterized as the heyday of the movement (cf. Andrews 1989, 402). Moreover, many of the most important contributions to the literature on subjective well-being continue to appear in the journal *Social Indicators Research*, which

[35] Another reason to think that people make global assessments of their lives, the authors suggest, is that they do not commit suicide more often (Andrews and Withey 1976, 64).

was established in 1974, and which is one of the most prominent journals of the movement (Andrews 1989, 402).

In the previous chapter we saw that research on subjective well-being as happiness appeared around the 1930's as part of the emergence of personality psychology. It was then that psychologists developed the confidence that they could measure personality traits and other features of the person, like happiness. However, it was the influence of the social indicator movement that cemented the idea that (average) subjective well-being could serve as a macro level indicator of quality of life, comparable to the standard economic indicators. Since in the view of the relevant authors the subjective measures were superior to economic or other objective ones, indicators of well-being emerged as a serious substitute for e.g. measures of GDP as indicators of the quality of life or well-being.[36]

2.2.9 Subjective well-being in economics

In the mid-1970's at least one economist was convinced that (average) subjective well-being could be used as a direct indicator of well-being (or welfare). The research reported in the above caught the attention of Richard A. Easterlin, whose paper 'Does Economic Growth Improve the Human Lot? Some empirical evidence' (1974) claims to be one of the first to offer a systematic examination of the "association between income and happiness" (Easterlin 1974, 90).[37] The reason why Easterlin, as an economist, is concerned with happiness is that he explicitly identifies it with welfare. He accepts

[36] According to Andrews (1989), the social indicator movement slowed down in the 1980's, partly as a result of funding cuts by the Reagan administration (Andrews 1989, 402). It would be interesting to explore how this fits with the fortunes of personality psychology, and with the evolution of the literature on subjective well-being. Is the rise of positive psychology associated with a comeback of the social indicators movement and personality psychology?

[37] He must mean in the economic literature, since psychologists had published on the topic for decades.

Pigou's distinction between social (or total) welfare and economic welfare (Easterlin 1974, 90). In Easterlin's view: "Happiness corresponds to the broader of the two concepts, that of social welfare, or welfare at large" (1974, 90). Unfortunately, he does not offer a more precise definition of happiness.[38] A central assumption behind the study is that the degree of happiness enjoyed by an individual is best determined by asking him or her. Easterlin writes: "Reliance is placed on the subjective evaluation of the respondent – in effect, each individual is considered to be the best judge of his own feelings" (1974, 92).

The data come from two different sources (Easterlin 1974, 91). The first come from studies in which a random sample of the population is asked a question of the Gurin et al. (1960) type: "In general, how happy would you say that you are – *very* happy, *fairly* happy, or *not very* happy?" (Easterlin 1974, 91). The second source is data from studies using the method devised by Cantril (1965), in which the subject is asked to identify where on "the ladder of life" he or she thinks she stands. The answers are coded on a scale from 0 to 10. The fact that these questions differ in various respects does not bother Easterlin. He writes: "Although the procedures differ ... the concept of happiness underlying them is essentially the same" (Easterlin 1974, 92). It is implicit that this common underlying concept of happiness is that same as Easterlin's own, so that both methods allow us to measure the degree of social (or total) welfare enjoyed by the subject.

To examine whether higher income is associated with greater happiness, Easterlin performs three different comparisons. First, he compares groups with different income within a country at a given time, and finds that happiness is indeed increasing in income. He writes: "There is a clear indication here that income and happiness are positively associated" (Easterlin 1974, 99). Second, he compares countries with different Gross National Product (GNP) per capita at a given point in time.

[38] In 2001, Easterlin writes: "Throughout this article, I use the terms happiness, subjective well-being, satisfaction, utility, well-being, and welfare interchangeably" (Easterlin 2001, 465).

In this case, he finds no clear association between wealth and happiness (Easterlin 1974, 105). Third, he compares the happiness of Americans at different points in time. He concludes: "As in the case of the international cross-sections, however, it seems safe to say that if income and happiness go together, it is not as obvious as in the within-country cross-sectional comparisons" (Easterlin 1974, 111).

This state of affairs raises the question of why a higher income is sometimes associated with greater happiness, but sometimes not. Easterlin asks: "Why do national comparisons among countries and over time show an association between income and happiness which is so much weaker than, if not inconsistent with, that shown by within-country comparisons?" (Easterlin 1974, 111). His answer is, in a nutshell, that people's happiness is not a function of a their absolute level of income, but of how their income compares to that of others in the same country. He writes:

> ... there is a "consumption norm" which exists in a given society at a given point in time, and which enters into the reference standard of virtually everyone. This provides a common point of reference in self-appraisals of well-being, leading those below the norm to feel less happy and those above the norm, more happy. Over time, this norm tends to rise with the general level of consumption, though the two are not necessarily on a one-to-one basis (Easterlin 1974, 112-113).

Easterlin recognizes that there are reasons for caution. He notes that the idea of happiness may not be present in all cultures (Easterlin 1974, 93). Moreover, the replies may be unduly influenced by temporary emotional states and therefore not be sufficiently stable (1974, 96). Also, people may not be capable of assessing their own emotional states, they may not be perfectly forthcoming, or they may be influenced by their perceptions of what is a socially appropriate response (1974, 96-97). The responses to questions may also be affected by other items in the same questionnaire, and by the wording of the question itself (1974, 98). Thus, there remains a real possibility that the results are biased (1974, 99).

In spite of these problems, Easterlin writes: "My own feeling is that while such bias may exist, it is not significant enough to invalidate the conclusions on the association between income and happiness" (1974, 99). In conclusion, he notes that "these conclusions raise serious questions about the goals and prospective efficacy of much social policy" (1974, 119). Easterlin considers the possibility that the net gain in happiness as a result of a general increase in income may be zero, but does not commit himself to that conclusion. Instead, he calls for more research into the topic (1974, 119).

This brief presentation of Easterlin's article illustrates that he adopted the central assumptions underlying psychological measures of well-being. First, the method he uses to measure welfare presupposes that welfare is a matter of a mental state, in this case happiness. Second, it is assumed that the individual herself is the best judge of her mental states. As a result, the best way to accurately estimate the degree of well-being enjoyed by an individual is to use her own judgment. Furthermore, it is taken for granted that subjects' responses can be quantified, and that they can be compared across individuals.

2.3 DISCUSSION

In the above, I have traced the historical roots of subjective measures of well-being. Far from a recent development, subjective measures can be traced back at least to the 1920's, and possibly earlier. The emergence of subjective measures can perhaps be ascribed to the 1918 article by Myerson, in which he promoted the field of eupathics, viz. the "more gracious sister" of eugenics, as he put it (Myerson 1918, 344). "Eupathics" – which translates to something like "well-feeling" – could be seen as just another name for "subjective well-being." It was in the late 1920's and early 1930's, however, that research into happiness really took off. This development can be attributed to

the studies by Davis (1929) and Hamilton (1929) into marital success, and the research by Watson (1930) in educational psychology. The development can also be described as an outgrowth of personality psychology, since this field had given psychologists the confidence that they could measures personality characteristics like happiness. In the 1960's and 70's, the push toward subjective measures was driven by developments in the epidemiology of mental health and gerontology, which presumably sped up the development of sophisticated sampling techniques necessary for large-scale research. Finally, the social indicators movement and its search for more direct measures of the quality of life appears to have been responsible for the wide-spread view that measures of subjective well-being – as estimated on the basis of responses by a representative sample of the population – can serve as macro level indicators of quality of life or well-being. Thus, a number of different endeavors have left their mark on the research project.

The history is not only longer, but also richer, than typically suggested. As we have seen, the researchers in the tradition discussed here had rather different purposes, used a wide variety of definitions of "well-being," "happiness," and "satisfaction," and invented a number of tools to measure it (some of which were ultimately rejected).[39] Even some of the very early studies used large samples – Davis (1929), for example, has a sample of 2200 – and longitudinal studies go back at least to Hart (1940).

In all essentials, the history explored here confirms the conclusions drawn by Porter (1995) in his general discussion of measurement in the social and behavioral sciences. First, Porter argues that the drive toward quantification emerged in applied rather than theoretical branches of social science, and that it did so not as a result of physics envy, but rather of a moral impulse to understand and improve a changing society (Porter 1995, viii; cf. Bateman 2001, 57; Levy 2001, 724). There is no doubt that subjective measures of well-being appeared in applied branches of

[39] See below for further discussion.

psychology and other fields. As we saw, subjective well-being first caught the attention of researchers interested in marital success and educational psychology, which on all accounts must be considered applied branches. At the time of writing, happiness and positive emotional states in general attract a good deal of attention from more theoretical branches of psychology and from neuroscience. See, for example, the collection *Well-Being* (Kahneman et al. 1999), which contains several articles e.g. on the neuroscientific basis for utility appraisals. Yet, the interest from the theoretical branches appears to be a more recent phenomenon.

There is little or no evidence that the proponents of subjective measures were trying to replicate the success of quantification in physics and the other hard sciences. At least, the literature reviewed in the above included no such suggestions. Rather, the main impetus behind the measures appears to have been a desire to better understand society in order to improve it. In the early literature, authors like Davis, Hamilton and Terman were clearly involved in trying to identify the factors that made for a successful marriage, with the understanding that a grasp of these factors would help them promote success in marriage. Similarly, authors like Watson were explicitly interested in finding out what makes (in particular) students happy, and he suggested that underlying his efforts were certain utilitarian leanings. The researchers in the tradition discussed were more much more impressed by fields like epidemiology – and its successful attempts at understanding and preventing disease – than by theoretical physics.

It is worth noting that – as far as I can tell – all the research surveyed so far is motivated by a belief that happiness (satisfaction, well-being) is something that should be promoted. Some – like Myerson (1918) – seem to suggest that happiness is the only thing that (ultimately) matters, and therefore the only thing that is (ultimately) worth promoting. Others do not go that far, and leave open the possibility that happiness is only one thing worth promoting among many. However, in

almost all cases it is unclear to what extent the authors consider happiness a good and how it relates to other goods (if there are other goods).

Second, Porter maintains that quantification often appeared as a means to overcome deficiencies in human judgment, and in particular a lack of trust and suspicions of arbitrariness (Porter 1995, 199; cf. Levy 2001, 724). Quite clearly, the researchers whose work has been reviewed in the above did not believe that unaided human judgment was sufficient to establish answers to questions about who is happy and why. They turned to empirical research to establish answers that could not be accused of being biased by personal, religious, and other considerations. Porter argues that it is no coincidence that the drive toward quantification was strongest in the weakest fields, such as psychology and social sciences, and in particular their applied subfields (Porter 1995, 200). In his view, practitioners in these fields were most insecure, and therefore most susceptible to outside and inside pressure. This idea is certainly consistent with the fact that the drive toward measuring happiness and the like appeared in fields like educational psychology.

Third, Porter suggests that the drive toward measurement was not impeded by the fact that researchers were unable to agree on the nature of the object under study (Porter 1995, 94-95; cf. Morgan 2001, 248). One of the fascinating aspects of the history reviewed here is how little agreement there is on the nature of well-being, happiness, satisfaction, and so on. Not only do various authors disagree among themselves regarding the proper understanding of these terms, but there are passages in which authors appear unable to agree with themselves. As for "happiness," Hart (1940) exhibits a certain degree of vacillation, when he offers two different definitions of the term. Both of his definitions have a distinct operationalist flavor. In contrast, Goldings (1954) defines happiness as a physiological state, though one that has consequences that are accessible through introspection. Many researchers in this tradition, however, simply omit to discuss what they mean by "happiness" and "satisfaction," though they appear to use the concepts interchangeably.

Some – like Burgess and Cottrell (1939) – just note that they use the concept in the "established" sense, and that they assume their subjects to do so too. In some cases – like that of Terman (1938) – they make it explicit that they hope to avoid philosophical discussion. When they do define the term, then, it is quite clear that different people have different accounts in mind.

As for satisfaction, we have seen that most early researchers appear to have used the term interchangeably with happiness. Although the questions given to subjects in the early studies invariably referred to happiness – of individuals or of marriages – the conclusions were often framed in terms of satisfaction as well as happiness. The underlying assumption appears to have been that "happiness" and "satisfaction" – at least for most purposes – can be treated as synonymous. Over time, psychologists started asking questions about satisfaction alongside questions of happiness, and they realized that answers to questions of satisfaction were not as highly correlated with answer to questions about happiness as they would have thought. As a result, they concluded that happiness and satisfaction really were two different traits, both of which were constitutive of well-being, and which had to be measured independently. Even so, few researchers have even tried to articulate adequate definitions of the term "satisfaction."

As we have seen, the lack of agreement regarding the nature of well-being did not stop people from continuing their efforts to measure it. The authors discussed above certainly proceeded in the absence of any such agreement. We have seen that psychologists since the 1930's have tried to use a variety of instruments to gauge the happiness of their subjects. There can hardly be said to be agreement on the proper way to measure subjective well-being. Some ask about happiness, others about satisfaction. Some use adjective batteries, some use bars and ladders, for the latter some use continuous scales, some use discrete scales, and so on. The time frame to which the questions refer is different too.

Nevertheless, at least during the episode surveyed in this chapter, there is more agreement on what measure of happiness to use than there is on the nature of happiness. To begin with, the choice of method appears to have been constrained since the beginning of what was in the process of becoming accepted practice in personality psychology. Moreover, over the time surveyed in this chapter there is a certain convergence in the tools used to gauge happiness. Thus, some of the many tools have completely dropped out of the picture. The notion that projection can be used to measure happiness, for example, appears to have been rejected. Also, there is a clear tendency to simply substitute a direct question – like that of Gurin et al. (1960) – for the more complex questionnaires (Diener et al. 1999, 277). For sure, the use of simple measures is probably in part a matter of convenience. Moreover, in order to attain more useful time series data it is critical to rely on the same tool as earlier researchers. However, many researchers have also drawn the conclusion that more complex tests do not radically increase the reliability or validity of the measure. At the end of the day, psychologists in this tradition appear to have concluded that self-revelation is as good as anything, though they obviously disagree about the time frame to ask about. It is true that according to the current consensus, "happiness" needs to be distinguished from "satisfaction," and both are constituents of well-being. At the same time, the correlation between answers to questions about happiness and satisfaction suggests to the researchers that for many purposes, a measure of either one can serve as a measure of both.

Incidentally, the historical survey sheds a good deal of light on the methods used by people who are active in the field of subjective well-being. It may seem that the methods used by the researchers interested in happiness and the like were grasped out of thin air. Indeed, this impression is sometimes encouraged by modern authors. Consider Frey and Stutzer, who simply state: "In general, it can be assumed that [individuals] are the best judges of when they are happy and when they are unhappy" (Frey and Stutzer 2002, 4). They support this assumption by referring to "a

sensible tradition in economics to rely on the judgment of the persons directly involved" (Frey and Stutzer 2002, 4). This move is particularly surprising, given that they – only a few pages earlier – attack the track record of economics in saying something useful about happiness (Frey and Stutzer 2002, vii). Anyway, this passage certainly gives the impression that psychologists and economists working on subjective well-being simply assume that people are reliable judges of their happiness. In reality, however, this is not just an unargued assumption, but a conclusion arrived at after a great deal of empirical research.

3.0 WELL-BEING IN PSYCHOLOGY AND ECONOMICS

Abstract. In this chapter, I explore what proponents of subjective measures mean by "well-being," that is, what account of well-being they use. My focus is on how the psychologists' account of well-being compares to that of proponents of traditional economic measures (e.g. income-based ones). My main thesis is that accounts of well-being differ radically both across and within disciplines. While economists tend to use preference-satisfaction accounts, psychologists almost without exception use mental state accounts; meanwhile, different members of each discipline adopt different kinds of accounts. I claim that this fact – which has not been properly acknowledged in the literature on subjective measures – has important implication for the assessment of these measures.

3.1 INTRODUCTION

The goal of this chapter is to explore what the defenders of subjective measures mean by "well-being." That is, the aim is to examine what account of well-being is used by these authors, who tend to be psychologists (though they also include a few economists). The account used by these authors is important, since it tells us what they purport to measure. In particular, I wish to compare the account of well-being used in the literature on subjective measures to that used in the traditional economic literature on welfare measurement. This literature – which, unsurprisingly, was written mainly by economists – tends to rely on traditional measures such as those based on real income. The fact that subjective measures of well-being are often presented as improvements over traditional economic measures of welfare may suggest to some that all these measures were designed to represent "the same thing," in the sense that psychologists use the same account of well-being as

economists. However, as we will see, economists and psychologists use radically different accounts of well-being.

The task of identifying underlying accounts of well-being is complicated by the fact that many psychologists and economists are remarkably unclear on what they mean by terms like "well-being" or "welfare."[40] In chapter 2.0, we saw that many authors would proceed from the assumption that the nature of e.g. happiness is commonly understood and widely agreed upon. Perhaps modern authors in the subjective well-being literature make a similar assumption. (Of course, even a cursory look at the philosophical literature on well-being would disabuse them of that notion.) Either way, it is often difficult to judge what account of well-being is used in a given text. As a result, evidence will have to be gathered from a variety of sources, and the resulting answers will have a certain provisional character.

My main thesis is that accounts of well-being differ radically across disciplines and across individuals. Relying on the taxonomy of Derek Parfit (1984), I argue that while economists tend to adopt preference-satisfaction accounts, psychologists almost without exception adopt mental state accounts. Meanwhile, there are important differences also within disciplines. Different economists appear to adopt different kinds of preference satisfaction account, whereas different psychologists seem to adopt different kinds of mental state accounts. Moreover, I argue these differences have been largely ignored in the literature, both by psychologists and economists, and that this fact helps explain why there has been relatively little useful communication and collaboration across disciplinary boundaries. Finally, I argue that the fact that different measures are based on different accounts of well-being has important implications for their assessment.

[40] The fact that some authors talk about measures of well-being when others talk about measures of welfare may seem to contradict the hypothesis that these measures are intended to reflect the same thing. This interpretation is undermined, however, by the evidence put forth in section 3.5.

In my view, clarifying the accounts of well-being used by psychologists and economists is worthwhile for a number of reasons. First of all, the discussion reveals the role played by purely philosophical assumptions – especially about the nature of well-being – in arguments for and against various measures of well-being. Moreover, increased clarity about the philosophical foundations of various measures should help us assess the arguments offered for and against these measures. Finally, it is possible that shedding light on philosophical foundations may lead to increased communication and collaboration across disciplines, something which should speed up the development of useful (and well understood) measures of well-being.

3.2 PHILOSOPHICAL ACCOUNTS OF WELL-BEING

As we saw in chapter 1.0, it is widely assumed that there is, as Scanlon (1993, 93) puts it, "a simple notion of individual well-being," interchangeably referred to as "well-being," "welfare," "quality of life," and so on. Over the years, philosophers have tried to shed light on the concept by developing and defending various accounts of well-being. The goal of this section is to offer an overview of the most popular philosophical accounts of well-being. Since it will be impossible to discuss each and every one in any detail, I rest content with identifying a number of different kinds of account, and discussing some of the accounts that fall into each class. The brief discussion is not intended to suggest that the accounts are free from problems.

The concept of well-being, as it is used here, needs to be sharply distinguished from the concept of financial well-being, or economic welfare, in the sense of access to economic resources (see Sen 1987, 16). While it is eminently plausible to assume that some economic resources are necessary for a life of well-being, as the term is used here, such resources are not constitutive of it.

The concept of well-being, as I use it, also needs to be distinguished from the concept of welfare as used e.g. in Nicholas Rescher's book *Welfare* (1972):

> On closer scrutiny, it emerges that welfare relates to the *basic requisites of a man's well-being* in general, but most prominently includes those basic concerns with health and economic adequacy to which we have become accustomed by such presently current terms as the "welfare state" or a "welfare worker" (Rescher 1972, 3-4, italics in original).

The concept of welfare as it is used in this passage is broader than that of economic welfare, as used above, since it has many dimensions of which economic welfare is but one (Rescher 1972, 4). On this view, happiness is the goal of welfare (Rescher 1972, 5).

For the purposes of this chapter, I divide accounts of well-being into three main classes: *mental state accounts*, *preference-satisfaction accounts*, and *objective list accounts*. Though different writers use different terms, this tri-partite division is standard. It is often traced to Derek Parfit (1984, 493-502) but it is also used in Griffin (1986, section 1), Daniel M. Hausman and Michael S. McPherson (1996, chapter 6), and Thomas M. Scanlon (1998, 99). Parfit (1984), for example, writes:

> What would be best for someone, or would be most in this person's interests, or would make this person's life go, for him, as well as possible? Answers to this question I call *theories about self-interest*. There are three kinds of theory. On *Hedonistic Theories*, what would be best for someone is what would make his life happiest. On *Desire-Fulfilment Theories*, what would be best for someone is what, throughout his life, would best fulfil his desires. On *Objective List Theories*, certain things are good or bad for us, whether or not we want to have the good things, or to avoid the bad things (Parfit 1984, 493, italics in original).

I discuss the three kinds of account in order. In spite of the fact that this taxonomy is widely used, it is not unproblematic. By the end of this section I will give an example of an account that does not seem to fit neatly into his categories.

3.2.1 Mental state accounts

According to one account, which has been immensely popular in the history of philosophy, well-being is a "mental state" or a "state of mind." What defines these accounts is, as Griffin sees it, that they all see welfare "as having to enter our experience" (Griffin 1986, 13). Griffin refers to this requirement as the *experience requirement* (Griffin 1986, 13). For this reason, Scanlon refers to mental state accounts as *experiential accounts* (Scanlon 1998, 99). He writes: "Experiential theories hold that the quality of life 'for the person who lives it' is completely determined by ... its experiential quality" (Scanlon 1998, 99), where "experiential quality" refers to "what it would be like to live it" (Scanlon 1998, 97). Similarly, on this view, "something contributes to well-being if, but only if, it affects the quality of one's experience" (Scanlon 1998, 100).

The most well-known mental state account, no doubt, is the view that Parfit refers to as *narrow hedonism* (Parfit 1984, 493). This view, which is associated with the classical utilitarians, equates well-being with happiness, and happiness with pleasure. On this account, as Raymond Plant (1991) writes:

> What is good for an individual is what tends to produce happiness or pleasure in the individual; what is bad or evil is what brings pain and unhappiness.... A person's interests, therefore, are what will help to maximise happiness or pleasure; what is against a person's interests is what will promote more unhappiness. Human well-being consists in the pursuit of happiness and human welfare consists in living a life with a preponderance of pleasure over pain (Plant 1991, 143).

In Griffin's words:

> Pleasure or happiness is presented as a 'state of feeling', and pain or unhappiness as a feeling on the same scale as, and the opposite of, pleasure or happiness. And the utilities of all our experiences are supposed to be determinable by measuring the amount of this homogeneous mental state that they contain (Griffin 1986, 7-8).

In brief, narrow hedonists believe that welfare is a homogeneous mental state, variously referred to as "happiness" or "pleasure."

A more sophisticated mental state account, associated with Henry Sidgwick, is that to which Parfit refers as *preference-hedonism*. Preference-hedonists agree that welfare is a matter of pleasure, but they reject the notion that pleasure is a homogeneous mental state. Parfit writes:

> What pains and pleasures have in common are their relations to our desires. On the use of 'pain' which has rational and moral significance, all pains are when experienced unwanted, and a pain is worse or greater the more it is unwanted. Similarly, all pleasures are when experienced wanted, and they are better or greater the more they are wanted. These are the claims of *Preference-Hedonism*. On this view, one of two experiences is more pleasant if it is preferred (Parfit 1984, 493, italics in original).

No matter our definition of pleasure, the important point is that according to preference hedonism, welfare is whatever mental state we desire. Scanlon puts the central idea as follows: "the experience of living a life is made better by the presence in it of those mental states, whatever they may be, which the person living the life wants to have, and is made worse by containing those states which that person would prefer to avoid" (Scanlon 1993, 186).

Preference hedonism is an example of what Griffin calls *eclectic accounts* of well-being, because they combine "a psychological element and a preference element" (Griffin 1986, 9). He writes: "'Utility' we could say, is 'desirable consciousness,' meaning by 'desirable' either consciousness that we actually desire or consciousness that we would desire if we knew what it would be like to have it" (Griffin 1986, 9). In spite of its eclectic nature, preference hedonism remains a mental state account because it satisfies the experience requirement: an event can affect a person's well-being if, but only if, it enters her experience.

3.2.2 Preference-satisfaction accounts

The second main class of accounts of well-being are referred to as *state of the world*, *desire fulfillment*, or *preference satisfaction* accounts. On such accounts, Griffin writes, "'utility' [is] states of the world which fulfil desires (e.g. economists' preference)" (Griffin 1986, 7). Such accounts do not require that a person who is well off experience any feelings of happiness or satisfaction. What they do require is that her desires are fulfilled (or that her preferences are satisfied), which does not come down to the same thing. Scanlon characterizes desire fulfillment accounts in the following way: "Desire theories reject the experience requirement and allow that a person's life can be made better and worse not only by changes in that person's states of consciousness but also by changes elsewhere in the world which fulfill that person's preferences" (Scanlon 1993, 186). Because these accounts reject the experience requirement, they imply that well-being "can, and it frequently does, come apart from any satisfaction or enjoyment. When you get what you want, you might like it, or you might not. You might not even know you've got it" (Moore and Crisp 1996, 599).

Preference satisfaction accounts differ depending on what preferences are supposed to matter to one's well-being. According to the most basic desire fulfillment account, what matters is the full range of desires or preferences that the agent actually has. In Griffin's words: "The simplest form of desire account says that utility is the fulfilment of *actual* desires. It is an influential account" (Griffin 1986, 10). Parfit puts it in the following way: "This [account] claims that what is best for someone is what would best fulfil *all* of his desires, throughout his life" (Parfit 1984, 494). Thus, people are well off to the degree that their actual preferences – that is, those preferences they actually happen to have – are satisfied.

Some philosophers have suggested that not all of our actual preferences are relevant to our well-being, and that the set of preferences that matter to our well-being should be restricted in some

way. For example, Parfit suggests that our welfare is affected only by the degree of satisfaction of our preferences about our own lives. He writes:

> Another theory appeals only to our desires about our own lives. I call this the *Success Theory*. This theory differs from Preference-Hedonism in only one way. The Success Theory appeals to *all* of our preferences about our own lives. A Preference-Hedonist appeals only to preferences about those features of our lives that introspectively discernible. Suppose that I strongly want not to be deceived by other people. On Preference-Hedonism it will be better for me if I believe that I am not being deceived. It will be irrelevant if my belief is false, since this makes no difference to my state of mind. On the Success Theory, it will be worse for me if my belief is false. I have a strong desire about my own life – that I should not be deceived in this way. It is bad for me if this desire is not fulfilled, even if I falsely believe that it is (Parfit 1984, 494).

Thus, preference hedonism and the success theory pull apart e.g. in cases when my preferences about my own life are satisfied and the agent in question is not aware of it.

Other philosophers have suggested that we restrict our attention only to those of our actual preferences that are fully informed, fully rational, or non-antisocial. Thus, we may want to restrict attention to informed preferences or desires. According to Scanlon, "informed desires are ones that are based on a full understanding of the nature of their objects and do not depend on any errors of reasoning" (Scanlon 1998, 101-102; cf. Griffin 1986, 11). Harsanyi writes that "we must exclude all clearly antisocial preferences, such as sadism, envy, resentment, and malice" (Harsanyi 1982, 56).

Some have suggested that the problems with the actual desire account should be dealt with by defining welfare in terms of *intrinsic* rather than informed preferences. Chappell and Crisp write:

> A simple desire theory fails immediately. I desire the glass of liquid, thinking it to be whiskey. In fact it is poison, so satisfying my desire will not make me better off. What desire theorists should say here is that it is the satisfaction of intrinsic preferences which counts for wellbeing. My intrinsic desire is for pleasure, the desire for the drink being merely derived (Chappell and Crisp 1998, 553).

Note that the intrinsic-desire account, as described here, appeals to the intrinsic desires people actually have, rather than to those that they would have under such-and-such counterfactual conditions.

A number of authors have also suggested that welfare is not a matter of the preferences that we actually have – restricted in some way or not – but of those preferences we would have under some suitable counterfactual conditions (Mongin and d'Aspremont 1998, 397). As Parfit puts it: "The obvious revision is to appeal not only to my actual preferences, in the alternative I choose, but also to the preferences that I would have if I had chosen otherwise" (Parfit 1984, 496). Thus, R. M. Hare (1982) argues that what matters in moral argument are those desires that the relevant individuals would have if they were perfectly prudent, "i.e. desired what they would desire if they were fully informed and unconfused" (Hare 1982, 28). As Sen and Williams (1982) point out, this implies that the preferences that count from the point of well-being may be very different from the preferences that the individuals actually have (Sen and Williams 1982, 10).

3.2.3 Objective accounts

The two kinds of account described so far can be referred to collectively as *subjective accounts*. In an article from 1975, Scanlon writes:

> By a *subjective criterion* I mean a criterion according to which the level of well-being enjoyed by a person in given material circumstances or the importance for that person of a given benefit or sacrifice is to be estimated by evaluating those material circumstances or that benefit or sacrifice solely from the point of view of that person's tastes and interests (Scanlon 1975, 656).[41]

[41] As is clear in this quote, Scanlon talks about *criteria for* rather than *definitions of* well-being. I am not sure if this difference has any notable implications for the present project.

Subjective accounts are so called because states of affairs are assessed by reference only to people's mental states, feelings, desires, preferences and the like. By contrast, we can identify what we may call *objective accounts*. Albert Weale (1998) writes: "Objective accounts of welfare appeal to the thought that there are features of the circumstances, position, of characteristics of persons that enable us to judge how well off they are" (Weale 1998, 704). Or, in Scanlon's words:

> By an *objective criterion* I mean a criterion that provides a basis for appraisal of a person's well-being which is independent of the person's tastes and interests, thus allowing for the possibility that such an appraisal could be correct even though it conflicted with the preferences of the individual in question, not only as he believes they are but even as they would be if rendered consistent, corrected for factual errors, etc. (Scanlon 1975, 658).[42]

Similarly, Chappell and Crisp claim, "there is now a return to ancient *ideal* theories of utility, according to which certain things are good or bad for beings, independently in at least in some cases of whether they are desired or whether they give rise to pleasurable experiences" (Chappell and Crisp 1998, 553). Such theories, as Parfit too points out, allow things to be good or bad for people *regardless* of what they want (Parfit 1984, 499). Those things may include, he says, "moral goodness, rational activity, the development of one's abilities, having children and being a good parent, knowledge, and the awareness of true beauty" (Parfit 1984, 499).

3.2.4 Well-being as multi-dimensional

By the end of his discussion, Parfit indicates that it may be possible to form a more plausible account of well-being by combining elements of different accounts. He says: "I shall end by mentioning another theory, which might be claimed to combine what is most plausible in these conflicting theories" (Parfit 1984, 501). He continues:

[42] Cf. Scanlon (1993, 188).

> We might ... claim that what is best for people is a composite. It is not just their being in the conscious states that they want to be in. Not is it just their having knowledge, engaging in rational activity, being aware of the true beauty, and the like.... On this view, each side in this disagreement saw only half the truth. Each put forward as sufficient something that was only necessary. Pleasure with many other kinds of object has no value. And, if they are entirely devoid of pleasure, there is no value in knowledge, rational activity, love, or the awareness of beauty. What is of value, or is good for someone, is to have both; to be engaged in these activities, and to be strongly wanting to be so engaged (Parfit 1984, 502).

Parfit's suggestion appears to have been picked up by Simon Keller, in a paper called 'Welfare and the Achievement of Goals' (2004). Keller writes:

> The idea here is that welfare is not a unitary concept like height or mass or monetary wealth; it is not the sort of thing that can be accurately represented by a single value. Welfare, rather, is like physical fitness. The ability to life heavy weights and the ability to run long distances, for example, are different, mutually irreducible aspects of fitness. The ability to run an extra mile always counts as an enhancement of physical fitness in one respect, but in one respect only. Sometimes it is outweighed by other aspects of fitness, so that one person is fitter than another even though the second person can run a mile further. And sometimes there's just no fact of the matter about whether the ability to run an extra mile contributes more or less than something else; if you can life ten kilos more than me and I can run a mile further than you, and if all else is equal, then you are fitter than me in one way and I am fitter than you in another – and there's no additional fact to be found about who is fitter "on the whole" (Keller 2004, 35).

In Keller's view, then, welfare has more than one dimension. He believes that you can be better off (or fitter) *simpliciter* than I am without scoring higher along all the relevant dimensions. However, he claims, there are also cases in which we cannot say that either one of us is better off (fitter) because there is no fact to the matter.

Although Keller does not (at last not in this paper) offer a complete account of welfare, he does suggest that one aspect of welfare is the achievement of goals. The paper defends what he calls

"The Unrestricted View" about the relationship between welfare and the achievement of goals.[43] According to this view: "An individual's achieving her goals *in itself* contributes to her welfare *regardless* of what those goals are" (Keller 2004, 28, italics in original). Thus:

> One aspect of an individual's welfare is her achieving her goals through her own efforts, regardless of what those goals are. It is not the only aspect of welfare, but it cannot be reduced to any of the others. Whenever an individual achieves a goal, she is better off in one respect, though she may be worse off in others (Keller 2004, 36).

As Keller points out, this implies that the achievement of goals makes a contribution to a person's welfare even if those goals are "crazy, self-destructive, irrational or immoral" (Keller 2004, 28-29).

This idea may seem closely related to the unrestricted desire-fulfillment account of welfare (Keller 2004, 28). However, Keller's proposal differs in several ways. First of all, of course, on Keller's accounts the achievement of goals is not the whole of welfare. There are other aspects to it that the achievement of goals. Second, Keller maintains, goals differ from mere preferences or desires. In his view, "to have something as a goal is, in part, to desire it, but you can desire something without having it as a goal" (Keller 2004, 32). Thus: "Taking something as a goal involves intending to put some effort into its achievement" (Keller 2004, 32). Third, achievements differ from mere attainments. The difference, in a nutshell, is the following: "To achieve a goal is to have its attainment be due in part to your own efforts" (Keller 2004, 33).

3.3 WELL-BEING IN ECONOMICS

In this section and the next, I will try to explore what accounts of well-being underlies various measures of well-being. Since subjective measures are often presented as alternatives to (and

[43] This view is, in fact, similar to the one endorsed by Scanlon (1998).

improvements over) traditional economic measures, it is appropriate to start out by saying a few words about the account of well-being used by economists. The concept of welfare is not incidental to economic theory and practice, as it has played a central role throughout the development of modern economics. For example, in his classic *The Economics of Welfare* A. C. Pigou identifies economic welfare with the subject matter of economics (Pigou 1960 [1920], 11). And as Hausman and McPherson put it, economists "typically rely on a theory in which the only normative concern is welfare and its distribution" (Hausman and McPherson 1997, 16).

It is worth noting also that the economist's interest in welfare has often been driven by a desire to formulate and evaluate policy. As Tibor Scitovsky writes: "Welfare economics supplies the economist – and the politician – with standards, at least with some standards, by which to appraise and on the basis of which to formulate policy" (Scitovsky 1951, 303). Commenting on the history of economics, Paul Samuelson (1947) writes: "Beginning as it did in the writings of philosophers, theologians, pamphleteers, special pleaders, and reformers, economics has always been concerned with problems of public policy and welfare" (Samuelson 1947, 203).

3.3.1 Welfare as preference satisfaction

Much of the economic literature is oddly silent on the question about the account of welfare or well-being underlying their efforts. Consider a graduate level textbook such as Mas-Colell et al. (1995). Although Mas-Colell et al. (1995) have a long section on welfare economics – "Part Five: Welfare Economics and Incentives" (Mas-Colell et al. 1995, 787-925) – they do not offer an explicit definition of "welfare." Nevertheless, there is little doubt that economists adopt some preference satisfaction account of well-being. That is, in the economic literature, agents are consistently assumed to be better off (in the sense of having more well-being) in state X than in state Y if and only if the bundle of goods in X is preferred to the bundle of goods in Y.

Consider first the use of (real) income, which remains is one of the most commonly used measures of welfare. Why would anybody think that a measure of real income represents somebody's welfare? Edward F. Denison illustrates the general line of thinking in the following passage:

> The output available to satisfy our wants and needs is one important determinant of welfare. Whatever want, need, or social problem engages our attention, we ordinarily can more easily find resources to deal with it when output is large and growing that when it is not (Denison 1971, 13).

The fundamental idea is that a higher income (or greater output) makes it possible for us to satisfy our wants and needs – that is, our preferences – to a greater degree. Denison adds: "The rationale is that, given the relative prices they face, people individually or collectively are free to spend their money in whatever way maximizes their satisfactions" (Denison 1971, 13). Thus, people can spend the resources available to them in such a way that they satisfy their preferences to the greatest extent possible.

Denison's line of thinking can be bolstered by reference to the theoretical result that, given a number of assumptions e.g. on individual preference and holding prices fixed, utility is strictly increasing in wealth. To show this Mas-Colell et al. (1995) define the indirect utility function $v(p,w) \equiv u(x(p,w))$, which tells us how much, utility the agent derives from $x(p,w)$, that is, from the bundle she would choose given p and w (Mas-Colell 1995, 56). Holding p fixed, the indirect utility function is strictly increasing in w (Proposition 3.D.3(ii) in Mas-Colell 1995, 56).[44] Since utility is used as an index of preference satisfaction, this means that increasing real income is associated with greater preference satisfaction. For real income to be a measure of *welfare*, however, economists add the identification of welfare and utility, i.e., the identification of welfare with preference satisfaction.

[44] A similar argument holds in the case of the preference approach.

This is what permits them to infer that the wealth level of the agent can be taken as a direct measure of the level of well-being enjoyed by the individual.

More sophisticated measures of welfare – such as those based on consumer surplus, compensating and equivalent variation, and the like – rely on very similar assumptions, and have welfare implications because the theoretical analysis is coupled with the assumption that welfare is a matter of preference satisfaction. Consider, for simplicity, the treatment of consumer surplus. Marshall defined consumer surplus of a good as "[the] excess of the price which [the consumer] would be willing to pay rather than go without the thing, over that which he actually does pay" (Marshall 1948, 124). Why should it matter what a person is willing to pay for a good? The idea is that a person's willingness-to-pay for a good reflects her *marginal valuation*, i.e. what the thing is worth to her on the margin.[45] This is the amount of money such that she would be indifferent between receiving one unit of the good and receiving the dollar amount for sure. The claim that individuals have a well-defined marginal valuation, and that their willingness-to-pay (as revealed by their market choices) reflects these valuations, are based on the principles of rational decision. Again, it can be shown that an increase in consumer surplus corresponds to an increase in utility (Johansson 1991, 41-42). Thus, as Per-Olov Johansson writes in his textbook in welfare economics: "One could say that consumer surplus expresses in observable monetary units an unobservable gain in utility; [by means of the notion of a consumer surplus] we transform the measurement problem from an unobservable dimension (utility) to an observable one (dollars)" (Johansson 1991, 41). Again, these claims have welfare implications because it is assumed that welfare is a matter of preference satisfaction.

[45] According to Hicks, "the marginal valuation of the unit" for a consumer is "the price at which he is just on the edge of purchasing" (Hicks 1943, 31).

The account of welfare implicit in economic analysis in general has been discussed by a number of scholars. Amartya Sen has pointed out the fact that the concept of preference in economic analysis does double duty. He writes:

> In economic analysis individual preferences seem to enter in two different roles: preferences come in as determinants of behaviour and they also come in as the basis of welfare measurements. For example, in the theory of general equilibrium the behaviour of individuals is assumed to be determined by their respective preference orderings.... At the same time, the optimality of an equilibrium, i.e. whether the market can lead to a position which yields maximal social welfare in some sense, is also examined in terms of preference with the convention that a preferred position involves a higher level of welfare of that individual. This dual link between choice and preference on the one hand and preference and welfare on the other is crucial to the normative aspects of general equilibrium theory. All the important results in this field depend on this relationship between behaviour and welfare through the intermediary of preference (Sen 1982, 66-67).

Thus, there are two central assumptions: first, that individuals act so as to maximize the degree to which their preferences are satisfied, and second, that individuals are well off to the extent that their preferences are satisfied. Taken together, the twin assumptions imply that people choose so as to maximize their welfare.

Other philosophers who have discussed the topic include Hausman, who writes: "Economists are implicitly saying that x is better than y for A if and only if A prefers x to y" (Hausman 1993, 180). In the words of Hausman and McPherson: "Welfare economics identifies welfare with the satisfaction of preferences. This identification is so automatic and ubiquitous that economists seldom realize how controversial it is" (Hausman and McPherson 1997, 17). If the identification really is automatic, this would help explain why authors like Mas-Colell et al. (1995) do not bother to articulate their account of welfare.

The claim that economists adopt some preference-satisfaction account is supported by the fact that some of them explicitly reject the experience requirement. That is, economists eschew the

notion that people necessarily derive a conscious feeling of satisfaction from goods that satisfy their preferences. As economists and philosophers Philippe Mongin and Claude d'Aspremont write:

> ... the utility of a thing or an action reflects the extent to which that thing or action is preferred to others, and has no meaning beyond that. Thus, the modern technical sense of "utility" [and therefore of "welfare"] not only excludes the commonsense notion of utility as usefulness, but also supersedes the old technical sense of utility as being related to pleasure and pain (Mongin and d'Aspremont 1998, 382).

Hausman and McPherson claim that "economists who have advocated a preference-satisfaction theory of well-being pride themselves on avoiding subjective notions about people's feelings" (Hausman and McPherson 1997, 17). As Mongin and d'Aspremont point out, the operative notions of utility and welfare are relatively recent. Pigou, for example, clearly adopted a mental state account, as when he wrote that "the elements of welfare are states of consciousness and, perhaps, of their relations" (Pigou 1960, 10).

3.3.2 Actual or ideal preferences?

While it is clear that economists adopt some preference-satisfaction account of well-being, it is less clear what preferences they believe should count. According to Mongin and d'Aspremont, economists rely on a preference-satisfaction account according to which what matters are actual preferences. They write:

> Following the most popular interpretation among 20th century writers, utility is a measure of actual preference satisfaction. "Actual" is meant to contrast the individual's preference underlying his behaviour with his rationally formed preferences. This interpretation underlies standard texts in economic theory, and pervades other social sciences as well as philosophy (Mongin and d'Aspremont 1998, 382).

The interpretation is strongly supported by the shape of the technical apparatus, which does not typically distinguish between actual and ideal preferences. In the vast majority of economic models,

there is but one preference ordering used for both the purposes identified by Sen (1982, 66-67; see section 3.3.1 above).

Some economists, however, explicitly recognize that people's actual preferences are based on false information and misguided inferences, and appear to reject the actual preference-satisfaction account. As Hausman and McPherson note:

> Economists recognize that this [theoretical world depicted in many standard economic models] is not the real world, and the fact that welfare is preference satisfaction in standard models does not imply that welfare is [actual] preference satisfaction in real life (Hausman and McPherson 1996, 73).

Though Hausman and McPherson do not name any economists in this passage, John C. Harsanyi is one of the economists who have addressed the issues head on. He writes: "To be sure ... a person may irrationally want something which is very 'bad for him'" (Harsanyi 1982, 55).

Harsanyi maintains that we need to distinguish between at least two different kinds of preferences. He writes: "In my opinion, social utility must be defined in terms of people's true preferences rather than in terms of their manifest preferences" (Harsanyi 1982, 55). Though he puts the point in terms of social rather than individual utility, it is clearly intended to cover both cases. Harsanyi defines *manifest* preferences as the agent's "actual preferences as manifested by his observed behaviour, including preferences possibly based on erroneous factual beliefs, or on careless logical analysis, or on strong emotions that at the moment greatly hinder rational choice" (Harsanyi 1982, 55). Manifest preferences contrast with *true* preferences, which are "the preferences he [the agent] *would* have if he had all the relevant factual information, always reasoned with the greatest possible care, and were in a state of mind most conducive to rational choice" (Harsanyi 1982, 55, italics in original). To determine what the true preferences of an individual are, we typically have to engage in some form of counter-factual reasoning. Harsanyi's account is clearly one according to which what

matters are the preferences that the agent in question would have under some counterfactual conditions.

Interestingly, even though he admits that actual and ideal preferences may fail to coincide, Harsanyi maintains that we should use manifest preferences as a basis for welfare measurement. He continues:

> But, while it is only natural to appeal from a person's irrational preferences to his underlying 'true' preferences, we must always use his own ['manifest'] preferences in some suitable way as our final criterion in judging what his real interests are and what is really good for him (Harsanyi 1982, 55-56).

The idea may be that for practical purposes, actual and ideal preferences are sufficiently similar that differences can be ignored. This view is likely not uncommon among economists who focus on ideal preferences. As Hausman and McPherson write, economists "often regard the differences between theory and reality as matters of detail" (Hausman and McPherson 1996, 73-74). Hausman and McPherson describe the idea (which they do not endorse) in the following way:

> Regardless of what human well-being truly is, the best *measure* of well-being is the extent to which [actual] preferences are satisfied (Hausman and McPherson 1996, 74, italics in original).

> Whatever welfare may be, there is no better *indicator* of welfare than people's preferences.... Sometimes people will want things that are harmful because of false beliefs, but who is likely to know better what is beneficial for an agent than the agent herself or himself? (Hausman and McPherson 1997, 17, italics in original).

There is another interpretation too, according to which Harsanyi is saying that it is morally imperative – perhaps for reasons of autonomy – to use the agent's own assessment of his or her welfare. This interpretation allows sharp divergences between actual and ideal preferences. Either way, there is evidence that economists disagree about the proper account of well-being. The main fault line appears to go between those who believe that what counts are actual preferences, and

those who believe that what counts are the preferences that the agent would have under specific counterfactual conditions, viz. if she were ideally rational and perfectly well-informed.

3.4 WELL-BEING IN PSYCHOLOGY

Unlike in economics, in psychology well-being has not always played a central role. In spite of the fact that what we now call subjective measures of well-being go back until the early decades of the twentieth century (see chapter 2.0), psychologists have not traditionally dedicated as much attention to well-being or welfare as economists have. With the rise of positive psychology, however, this is beginning to change (Seligman and Csikszentmihalyi 2000, 5). In this section, I discuss the account of well-being that is implicit or explicit in the work of psychologists and the occasional economist who defend subjective accounts of well-being.

3.4.1 Well-being as a mental state

There is little doubt that the psychologists use some kind of mental state account of well-being. This is clear, among other things, from their adherence to the experience requirement. In the literature on subjective measures, well-being is often described as a matter exclusively of individual subjective, hedonic, or affective experience. For example, David G. Myers quotes Madame de la Fayette as saying: "If one thinks that one is happy, that is enough to be happy," and adds that "like Madame de La Fayette, social scientists view well-being as a state of mind. Well-being, sometimes called 'subjective well-being' to emphasize the point, is a pervasive sense that life is good" (Myers 1992, 23; cf. 1992, 27). Myers evidently takes well-being to be something thoroughly subjective; the explicit

reference to states of mind strongly suggests that what he has in mind is some type of mental state account.

In his 1984 review, Ed Diener offers a useful discussion about the notion of well-being as it is used in the psychological literature. He writes:

> The area of subjective well-being [SWB] ... is subjective.... [It] resides within the experience of the individual. Notably absent from definitions of SWB are necessary objective conditions such as health, comfort, virtue, or wealth. Although such conditions are seen as potential influences on SWB, they are not seen as an inherent and necessary part of it. (Diener 1984, 543)[46]

This quote confirms that, in Diener's field, an individual's well-being is defined not by the objective circumstances in which she finds herself, but by her subjective experience. Diener and Suh (1997) reinforce this point:

> Subjective well-being research ... is concerned with individuals' subjective experiences of their lives. The underlying assumption is that well-being can be defined by people's conscious experiences – in terms of hedonic feelings or cognitive satisfactions. The field is built on the presumption that to understand the individuals' experiential quality of well-being, it is appropriate to directly examine how a person feels about life in the context of his or her own standards (Diener and Suh 1997, 191).

Diener and Suh, like Myers and the scientists to whom he refers, evidently use the term "well-being" interchangeably with "subjective well-being." The fact that Diener and Suh argue that well-being is not only *concerned* with individual's subjective experiences, but is *defined* by them, strongly suggests that Diener and Suh adhere to the experience requirement.

Several authors emphasize the subjective character of well-being, as they use the term, by contrasting it with what could be called "objective measures of well-being," that is, e.g. social and economic indicators. Diener makes this point in the 1984 quote above, where he says that "objective

[46] References have been omitted.

conditions such as health, comfort, virtue, or wealth" may affect, but do not constitute, well-being in his sense (Diener 1984, 543). Kahneman adds: "Objective happiness is not to be confused with good fortune, which is an assessment of the circumstances of someone's life" (Kahneman 1999, 5). It is possible to enjoy good fortune, in Kahneman's view, without enjoying objective happiness. He does not say what he means by good fortune, but presumably refers to income, health, and so on.

Angus Campbell (1976), quoted in Diener (1984, 543), expands on this point. Campbell writes:

> The gross national product, important as it undoubtedly is, is clearly not the ultimate touchstone against which the quantum of happiness in this country can be assessed.... We now have a growing array of national statistics describing noneconomic aspects of American life. Nearly all of these social indicators describe events, behaviors, or characteristics of individuals that are reported through governmental institutions of one sort or another and do not depend on the individual's description of his own life. They are what might be called *objective* indicators (Campbell 1976, 117-118, italics in original).

Later on, he continues:

> If we are primarily concerned with describing the quality of life experience of the population, we will need measures different from those that are used to describe the objective circumstances in which people live. We will have to develop measures that go directly to the experience itself. These *subjective* measures will surely not have the precision of indicators that are expressed in number of dollars, units of time, or numbers of square feet, but they will have the great advantage of dealing directly with what it is we want to know, the individual's sense of well-being (Campbell 1976, 118, italics in original).

Quality of life, on Campbell's view, is a function of individuals' *sense* of well-being. If we want to study the quality of life, we need to study the individuals' *experience* of their lives, not the objective characteristics of their existence.

The mention of individuals' "sense of well-being" allows at least two different interpretations. There is a conceptual distinction between individuals' actual well-being (how well off

they really are), their perceived well-being (how well off they think they are), and their revealed well-being (how well off they say they are). It is not quite clear from Campbell's quote whether "sense of well-being" should be understood as their actual or perceived well-being. Presumably, Campbell (and other psychologists who use similar locutions) believe that well-being is distinct from perceived well-being, so that it is conceptually possible for individuals to be mistaken about how well off they are. The reason why Campbell and others do not discuss these distinction in any great detail, though, in all likelihood is that they believe that people (by and large) know how well off they are, and when prompted truthfully reveal how well off they think they are.[47] Yet, this does not change the fact that there is a conceptual distinction to be drawn.

3.4.2 What kind of mental state account?

Because of their fairly evident adherence to the experience requirement, it seems fair to assume that the psychologists who work on subjective measures adopt some kind of a mental state account. It is more difficult to say, however, what precise account they adopt. Many psychologists appear committed to what Parfit calls narrow hedonism, according to which well-being is identified with happiness, and happiness is associated with pleasure. According to Martin E. P. Seligman and Mihaly Csikszentmihalyi: "Subjective well-being is a more scientific-sounding term for what people usually mean by happiness" (Seligman and Csikszentmihalyi 2000, 9). It is not obvious what Seligman and Csikszentmihalyi think people "usually" mean by happiness, but presumably what they have in mind is a simpler account, such as narrow hedonism, rather than a more sophisticated one, such as preference hedonism.

[47] Problems associated with the measurement of well-being will be discussed in chapter 4.0.

Kahneman's writings give a somewhat better idea of what he and his co-authors may have in mind. In "Objective Happiness" (Kahneman 1999), lead essay in the collection, "well-being" is simply used as a synonym for "happiness." He writes: "We distinguish two notions of happiness, or well-being (the two terms are used interchangeably in this chapter)" (Kahneman 1999, 5). Next, Kahneman draws a distinction between subjective and objective well-being. He writes: "*Subjective happiness* is assessed by asking respondents to state how happy they are. *Objective happiness* is derived from a record of instant utility over the relevant period" (Kahneman 1999, 5, italics in original). As this quote suggests, subjective happiness (or well-being), in Kahneman's terms, is a matter of how one feels at a given moment in time. Objective happiness is derived from a continuous record of one's subjective happiness. Kahneman writes: "the objective happiness of Helen in March should be measured by the average of the instant utility that she experienced during that period, after appropriate rescaling" (Kahneman 1999, 6).[48] If I understand Kahneman correctly, instant utility is identical to subjective happiness. Kahneman writes: "Being pleased or distressed is an attribute of experience at a particular moment. I will label this attribute *instant utility*" (Kahneman 1999, 4). The use of the terms "subjective" and "objective" must be considered unfortunate, especially the social indicator approach, which studies unemployment, infant mortality, and so on as "objective measures of well-being" (see section 1.2.3). To clarify, Kahneman adds: "Objective happiness, of course, is ultimately based on subjective data" (Kahneman 1999, 5). Kahneman's account is unique in its explicitly dynamic character. Yet, the remarks quoted here – especially his references to being pleased or distressed – also point in the direction of narrow hedonism. What matters, to Kahneman,

[48] Elsewhere he suggests that somebody's objective happiness is the time-integral of his or her subjective happiness, after appropriate rescaling (e.g. Kahneman (1999, 6). By the way, the idea of "appropriate rescaling" is one that I will ignore in the present context.

is what it feels like to be in a certain state (not e.g. what the relationship between the state and the individual's preferences as the preference-hedonist would have it).

Another understanding of well-being appears to have developed after psychologists started rejecting the synonymy of "happiness" and "satisfaction." After realizing that answers to questions about happiness and satisfaction do not correlate as highly as they had expected, psychologists started talking about happiness and satisfaction as separate (e.g. Campbell et al. 1976). Thus, authors like Campbell et al. (1976) argued that satisfaction is a "judgmental or cognitive experience" whereas happiness is an "experience of feeling or affect" (cf. Campbell et al. 1976, 8). While it would be possible to identify well-being with either happiness or satisfaction, it appears that many psychological researchers adopted the view that well-being is a matter of both. This approach appears most consistent with a view according to which there are two homogeneous, irreducible, mental states, happiness and satisfaction, both of which are constitutive of well-being.

As a result of many psychologists' emphasis on satisfaction, they cannot accurately be described as narrow hedonists. It would be possible for them to adopt an account according to which happiness is has a dual character: an affective component referred to as pleasure (or some such) and a cognitive component referred to as satisfaction. Such an account may be referred to as *dual hedonism*. However, such a characterization would be slightly misleading, because of many psychologists' emphasis on happiness and satisfaction as separate components of well-being. Thus, we could call them *dual state* accounts, in order to emphasize the fact that well-being is supposed to consist of two irreducible states without prejudging the issue of the definition of happiness. Incidentally, a similar analysis can be given for those psychologists who insist that well-being is constituted by three irreducible components: positive affect, negative affect, and satisfaction (following the work of Bradburn and Caplovitz 1965; cf. section 2.2.5). Call such an account *triple state* account.

Other psychologists suggest that well-being is a great more complex. Thus, Diener and Seligman (2004) write: "Well-being, which we define as people's positive evaluations of their lives, includes positive emotion, engagement, satisfaction, and meaning" (Diener and Seligman 2004, 1). This definition, if it is one, still appears to satisfy the experience requirement, assuming that positive evaluations (whatever they are) in fact enter people's experience. If so, Diener and Seligman do adopt a bona fide mental state account. It is less clear how to understand the exact variety of mental state account presupposed in this passage. Perhaps the idea is that all mental states that satisfy the description "positive evaluation" – whatever they happen to be – are jointly constitutive of well-being. If so, perhaps the underlying account of well-being could be referred to as *multiple state* accounts.

There is also a way in which psychologists could be understood as presupposing what Griffin refers to as an eclectic account of well-being (see section 3.2.1). Assuming that people in fact desire the two mental states happiness and satisfaction, what matters to well-being (from the point of view of certain eclectic accounts) is the degree to which people have these mental states. Assuming that people desire those mental states that can be described as "positive evaluations," what matters to well-being (from the point of view of certain eclectic accounts) are those mental states that can be described as positive evaluations. The psychologists could perhaps be described as preference hedonists, although the psychologists' understanding of the concept of happiness appears to differ from that of the preference hedonists as described by Parfit. The psychologists' understanding of "happiness" seems closer to that of the narrow hedonists, in fact.

In yet other passages, psychologists leave open the possibility that well-being may have constituents that are not mental states. Consider the following passage, from Kahneman (2000a):

> Objective happiness is not proposed as a comprehensive concept of human well-being, but only as a significant constituent of it. Maximizing the time spent on the right side of the affect grid is not the most significant value in

> life, and adopting this criterion as a guide to life may be morally wrong and perhaps self-defeating as well. However, the proposition that the right side of the grid is a more desirable place to be is not particularly controversial.... Objective happiness is a common element of many conceptions of well-being (Kahneman 2000a, 691).

Kahneman's suggestion that the pursuit of happiness may be immoral and self-defeating is intriguing. It is unfortunate that Kahneman does not say more about it, or about what the other constituents or elements of well-being may be. Either way, however, there is a certain amount of tension here, between Kahneman's (2000a) thought that pursuing happiness may be morally wrong, and his insistence that public policy should be designed so as to maximize objective happiness. Consider the following quote: "In the present framework ... it is objective happiness that matters. Policies that improve the frequencies of good experiences and reduce the incidence of bad ones should be pursued" (Kahneman 1999, 15; also quoted in section 1.3). Notice that this claim, as it stands, is a strong one. It would have been much weaker if he had added "all things equal," or some such.

3.4.3 Summary

In the previous sections, I have argued that both psychologists and economists are remarkably unclear on what they mean by "well-being," but that accounts of well-being differ both across and within fields. While economists tend to adopt preference satisfaction accounts of well-being, psychologists almost exclusively adopt mental state accounts. Meanwhile, economists appear to disagree on what preferences matter to well-being. While some appear to presuppose that what matters are actual preferences, others specify that they think ideal preferences are what matters. Meanwhile, some psychologists seem to adopt narrow hedonism, whereas others use dual or multiple state accounts. They could also be understood as adopting preference hedonism or some other eclectic account.

3.5 THE LACK OF APPRECIATION FOR DIFFERENCES ACROSS FIELDS

In this section, I will argue that the fact that accounts of well-being differ dramatically across and within fields has been largely ignored in the literature. I will discuss the work of several prominent contributors to the literature on subjective measures, and show that they have failed to properly acknowledge the fact that they adopt radically different accounts of well-being.[49] The result is a misrepresentation of the economists' account of well-being, as well as their account of rationality, and great deal of confusion when it comes to the alleged advantages of the subjective measures.

Kahneman, Diener and other prominent proponents of subjective measures often criticize traditional economic measures for being "indirect." They do not, however, properly acknowledge that those measures are not intended to reflect well-being as psychologists see it. Thus, Kahneman, Diener and Schwarz (1999) write that "economics assesses variables that are only indirect indicators of something else – of subjective fulfillment" (Kahneman et al. 1999, xii). As we saw in section 3.3.1 above, modern economists see well-being as a matter of preference satisfaction, and they reject the notion that the satisfaction of preferences need to be associated with any feelings of satisfaction. Thus, they certainly do not see real income as an indicator of subjective fulfillment.

In their review "Beyond Money: Toward an economy of well-being," Diener and Seligman (2004) argue:

> Although economics currently plays a central role in policy decisions because it is assumed that money increases well-being, we propose that well-being needs to be assessed more directly, because there are distressingly large, measurable slippages between economic indicators and well-being (Diener and Seligman 2004, 1).

[49] In this section, I focus on the psychologists' critique of economic measures. In the next section, I will discuss economists' critique of subjective measures. The issues remain largely the same.

If the term "well-being" in this passage is interpreted in the manner of the psychologists, the first claim is obviously false: economic measures are used because it is assumed that they reflect degrees of preference satisfactions, not degrees of positive mental states. However, if the term is used as the economists understand it, Diener and Seligman have no support for the second claim – that there are slippages between economic indicators and well-being – since the evidence they report relates to well-being understood in terms of positive mental states (Diener and Seligman 2004). Thus, it appears that the authors equivocate on the meaning of "well-being."

Again, after noting that economics "reigns unchallenged in the policy arena, as well as in media coverage of quality-of-life indicators," Diener and Seligman (2004, 2) write:

> Money, however, is a means to an end, and that end is well-being. But money is an indirect surrogate for well-being, and the more prosperous a society becomes, the more inexact a surrogate income becomes. The measurement of well-being has advanced sufficiently that it is time to grant a privileged place to people's well-being in policy debates, a place at least on a par with monetary concerns. After all, if economic and other policies are important because they will in the end increase well-being, why not assess well-being more directly? (Diener and Seligman 2004, 2).

On the assumption that Diener and Seligman use the term "well-being" in the psychologists' sense, it fails to accurately represent the reason why economic measures are taken to be important.

Some passages fail to properly acknowledge the economists' account of rationality as well as their account of well-being. Kahneman et al. (2004) write:

> Economists have traditionally eschewed direct measures of well-being on methodological grounds: the private nature of experience and the discomfort of making interpersonal comparisons. Instead income is often used as a proxy for opportunities and well-being. If people are not fully rational, however, their choices will not necessarily maximize their experienced utility, and increasing their opportunities will not necessarily make them better off.

> Direct measures of experienced utility become particularly relevant in a context of bounded rationality (Kahneman et al. 2004, 429).[50]

When Kahneman claims that violations of rationality mean that "choices will not necessarily maximize experienced utility," this suggests – though it does not in fact logically entail – the claim that rational behavior amounts to behavior that maximizes experienced utility. As we saw above, however, the neo-classical view is that rational behavior amounts to behavior that maximizes preference satisfaction, which is not to say that it maximizes experienced happiness. Again, preference satisfaction needs to be distinguished from feelings of satisfaction, pleasure or happiness.

It may be objected that the psychologists make it clear that their concept of well-being is different from that of the economists by occasionally using expressions like "subjective well-being," as in "measures of subjective well-being" (as opposed to "subjective measures of well-being"). I should say that I have no objection to the use of the term "subjective well-being" to denote whatever the psychologists have in mind, so long as it is clearly defined and distinguished from well-being *tout court*. However, this is but a weak objection to the argument offered here. As I hope to have shown in this section, there remain a great number of passages in which psychologists do not make it sufficiently clear that their understanding of "well-being" differs radically from that of the economists. Indeed, there are passages in which there is a certain amount of equivocation. Thus, it is fair to say that the psychologists have not properly acknowledged this difference.

In my view, the quotes presented in this section show that psychologists have failed to properly acknowledge that economists and psychologists use different accounts of well-being. This failure, I would argue, has several important implications. First of all, these passages end up misrepresenting the account of well-being adopted by the economists (and perhaps their account of rationality as well); economists do not believe that their measures reflect subjectively experienced

[50] References have been omitted.

states, and they do not believe that rational action maximizes "experienced utility" in Kahneman's sense. Second, and consequently, the quotes misrepresent the economists' case in favor of traditional measures; this case is not based on any proposition related to subjectively experienced mental states, as the passages suggest. While the economists' case can be, has been, and indeed should be, criticized, it cannot be properly assessed until it has been properly articulated. Third, the passages quoted in this section suggest, erroneously, that subjective measures were designed to represent the same account of well-being as economic measures. Finally, these passages obscure the fact that psychologists' case in favor of subjective measures depends essentially on a particular – and, as we will see shortly, implausible – philosophical account of well-being.

3.6 WHY DO ACCOUNTS OF WELL-BEING MATTER?

In this section I discuss why it matters that psychologists and economists use different accounts of well-being. In brief, the main reason (as I see it) is that arguments for and against measures of well-being typically – perhaps always – proceed from some assumptions about the nature of well-being. This is equally true, it appears, for psychologists and economists. As a result, the plausibility of psychologists' and economists' arguments is irrevocably tied to the plausibility of their underlying account of well-being. This is not to say, however, that a given measure of well-being is plausible only if the account of well-being used by its proponents is correct. It may well be possible to defend a given measure of well-being on several different grounds.[51]

In chapter 1.0, we saw that the psychologists' case for the use of subjective measures for purposes of public policy relies heavily on the proposition that well-being (as the psychologists

[51] See section 3.7 for a more detailed discussion of this point.

understand it) has some privileged normative status. Recall, for example, how Angus Campbell argued that happiness has "almost universal currency as a recognized, if not uniquely important, component of the quality of life experience" (Campbell 1976, 119), and how Andrew Oswald added: "Economic things matter only in so far as they make people happier" (Oswald 1997, 1815). The reason, in this view, why we should care about well-being as happiness is that happiness is what ultimately *matters*. Similarly, the psychologists' argument against economic measures as unsuitably indirect (see section 3.5 above) proceeds from the assumption that well-being is some mental state or other, and that there is a great deal of "slippage" between real income and whatever mental state constitutes well-being. If it turned out that mental state accounts of well-being were mistaken, this would be a serious blow to the arguments offered by the psychologists.

Something very similar is true in the case of the economists. The arguments why indicators of real income, consumer surplus, and equivalent and compensating variation are supposed to reflect well-being or welfare is that they are an index of preference satisfaction. Textbooks like Mas-Colell et al. (1995) defend the use of such measures as measures of well-being on the basis that they can be shown to be utility functions, and utility is understood as no more than an index of preference satisfaction (see section 3.3.1 above). If it turned out that preference satisfaction accounts of well-being were mistaken, this would be a serious blow to the arguments offered by the economists.

The fact that psychologists and economists adopt such different accounts of well-being, in combination with the fact they do not seem to acknowledge these differences – could help explain why there is relatively little communication and collaboration across disciplinary boundaries. It is true that such interactions have become slightly more common in recent years, as more economists – like Frey and Stutzer (e.g. 2002) – have come to adopt the psychologists' definition of "well-being." Yet, there is little fruitful exchange between economists who adopt more traditional approaches, and psychologists and economists who have committed themselves to the subjective

measures. Insofar as proponents of a given measures fail to appreciate that their arguments proceed from different premisses about the nature of well-being, relative to those of their opponents, their arguments may fail to be convincing.

In the rest of this section, I offer a real-life illustration of the manner in which arguments for and against different measures of well-being depend on claims regarding the nature of well-being. In passing, the example will also illustrate the importance of the choice of welfare measure for purposes of the design and evaluation of public policy. The discussion appears in the work of the economist Anders Åslund.[52] Åslund served as an advisor to the Russian government during the post-Soviet reforms during the early 1990's. He was one of the main proponents of shock therapy, which was a policy of rapid privatization and deregulation with an eye to quickly transforming Russia into a modern, market-based democracy. According to e.g. World Bank (2004) figures, the subsequent development was disastrous. GDP in constant US dollars dropped by almost 35 percent between 1991 and 1995, and life expectancy for men dropped by almost five years (cf. Angner in progress). Åslund, however, has argued that the policy was a great success. According to his own calculations, GDP decreased by no more than 6 percent during the same period (Åslund 2001, 15).

Given Åslund's relatively cheerful figures, it is something of a paradox that Russians consistently report that they are less happy after the reforms than they were before. In his paper 'Russia's Success Story,' Åslund asks "why are Russians not happier?" (Åslund 1994, 66). He replies: "The main answers are that many people do not realize that they are not worse off and that they care about things other than their material well-being" (Åslund 1994, 66). Thus, Åslund appears to draw a sharp distinction between happiness and well-being. He does not dispute the claim that Russians are less happy than they used to be, but he does dispute the claim that they are worse off. In this view, then, it appears that being unhappy has few or no implications for one's well-being.

[52] See Angner (in progress) for an extended discussion about Åslund's role in the Russian reforms.

How then is it possible for people to be unaware of the fact that they are better off? Åslund suggests an answer to this question in a 2001 paper, in which he asks "[why] people indicate in opinion polls that the material situation has deteriorated?" (Åslund 2001, 17). Part of the answer, according to Åslund, is that "people do not think about whether total welfare rises or falls (Pareto optimality)" (Åslund 2001, 17). The mention of "Pareto optimality" is interesting, because the concept is defined in terms of preference satisfaction. Åslund adds: "Public sentiment about the general situation should thus be taken with a great deal of skepticism" (Åslund 2001, 17). Again, the problem seems to be that people in general have a mistaken idea about the nature of welfare, in that they do not think about it in terms of Pareto optimality and preference satisfaction.

Since Åslund himself is not terribly explicit, there is room for interpretation. Nevertheless, it seems clear that he wants to divorce questions of happiness from questions of welfare, so that one's happiness carries few or no implications for one's well-being. The reason, I take it, is that people are unhappy as a result of having the wrong idea about their own well-being. This explains why their happiness (and their perceptions of their well-being) should carry little weight. Clearly, what distinguishes Åslund from his critics is the account of well-being that they adopt. Presumably, Åslund uses some preference satisfaction account (the concept of Pareto optimality is defined in terms of preferences), whereas his critics adopt some view according to which people's subjectively experienced mental states are directly material to their well-being.

In the context of Åslund's argument, questions about the correct account of well-being become critical. It is hard to imagine that the disagreement between Åslund and his critics can be settled in the absence of some agreement on the appropriate account of well-being. Thus, the example illustrates the importance of accounts of well-being to the choice of measure of well-being. The example also illustrates how important questions about the measurement of well-being can be.

In the case of post-Soviet Russian reforms, it appears that the judgment regarding whether the reforms were a success or not hinge, to a very great degree, on the choice of measures of well-being.

3.7 THE PLAUSIBILITY OF PSYCHOLOGISTS' ACCOUNTS OF WELL-BEING

If the argument in the previous section is correct, the case in favor of subjective measures of well-being depends to some significant degree on the plausibility of the underlying account of well-being. Thus, it is worth considering the plausibility of the accounts adopted by the psychologists (while keeping in mind that different psychologists adopt slightly different accounts of well-being). My purpose in this section is to review some arguments to the effect that those accounts in fact are implausible. The purpose is not to settle the question about the correct account of well-being; doing so would be far too large a project. Neither do I aspire to develop an original argument to that effect. In the main, I will point to a long philosophical tradition that considers mental state accounts like those used by proponents of subjective measures utterly implausible. The arguments against them are be so strong that, to the best of my knowledge, no contemporary philosopher is willing to defend them. In the case of hedonism, Sumner (1996) puts it as follows:

> Time and philosophical fashion have not been kind to hedonism. Although hedonistic theories of various sorts flourished for three centuries or so in the congenial empiricist habitat, they have all but disappeared from the scene. Do they now merit even passing attention, for other than nostalgic reasons? (Sumner 1996, 83).[53]

The fact that no living philosopher appears willing to defend a particular position must be considered unusual; no matter how implausible a point of view, there are typically philosophers willing to accept it.

[53] See below, though, for more about Sumner's views on well-being.

Griffin has summarized what he sees as the main deficiencies of narrow hedonism and related accounts:

> The trouble with thinking of utility 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. What one mental state runs through them all in virtue of which we rank them as they do? Think of the following case. At the very end of his life, Freud, ill and in pain, refused drugs except aspirin. 'I prefer,' he said, 'to think in torment than not to be able to think clearly.' But can we find a single feeling or mental state present in both of Freud's option in virtue of which he ranked them as he did? (Griffin 1986, 8; cf. Parfit 1984, 493).[54]

Griffin's point, as I understand it, is that there is no one mental state such that the definition of well-being as that mental state is not too narrow or too broad. Now, Griffin's claim is a quite general one, but if restricted to a single mental state like happiness or satisfaction (as the psychologists appear to understand it) the claim is undoubtedly correct. It is easy to imagine cases of people (like Freud as described in the quote above) who are well off in spite of the fact that they are unhappy, or happy in spite of the fact that they are not well off. The same objection, mutatis mutandis, can be raised against dual and multiple state accounts, given the way the psychologists understand happiness, satisfaction, and the like. It is quite possible to think about people who are neither happy nor satisfied, in the psychologists' sense, but who can nevertheless correctly be described as well off.

One modern philosopher who is well aware of these objections, and who nevertheless defends a kind of a mental state account, is Sumner (1996). Sumner is particularly interesting in the present context because his discussion of happiness and satisfaction is inspired by Campbell and others (cf. Sumner 1996, 149 *idem*). However, Sumner's account involves an *authenticity requirement*. In his view, welfare consists in authentic happiness, viz. "the happiness of an informed and autonomous subject" (Sumner 1996, 172). This constitutes a rather large departure from accounts

[54] A reference has been omitted.

that do not have such a requirement, and there is no evidence that the psychologists discussed in the above accept such a requirement. If they did, they would have to present reasons why measures apparently designed to reflect happiness *tout court* should also reflect authentic happiness. Surely, when people say that they are happy, often what they are reporting is (on Sumner's account) inauthentic happiness. Thus, even if Sumner turned out to be correct, this would not imply that the psychologists were right about the nature of happiness.

Although it may seem that preference hedonism – according to which well-being is a matter of *desired* mental states – constitutes a more plausible alternative than more old-fashioned mental state accounts, in fact it does not fare much better. Again, there seems to be cases in which people can be well off even if they are not in a desired mental state, and they can be in a desired mental state without being well off. As Griffin puts it: "The trouble with this eclectic account is that we do seem to desire things other than states of mind, even independently of the states of mind they produce" (Griffin 1986, 9). This is the same point, I believe, that Robert Nozick (1974) makes in a particularly fanciful manner. Nozick asks us to consider hooking ourselves up to a machine that can offer us any experience we would like. In his words:

> Suppose there were an experience machine that would give you any experience you desired. Superduper neuropscyhologists could simulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book.... Would you plug in? *What else can matter to us, other than how our lives feel from the inside?* (Nozick 1974, 42-43, italics in original).

What Nozick wants to say, of course, is that our well-being depends not only on how things "feel from the inside." Griffin suggests that we do not need fanciful science-fiction to make the point:

> I prefer, in important areas of my life, bitter truth to comfortable delusion. Even if I were surrounded by consummate actors able to give me sweet simulacra of love and affection, I should prefer the relatively bitter diet of their authentic reactions. And I should prefer it not because it would be morally better, or aesthetically better, or more noble, but because it would make for a better life for me to live (Griffin 1986, 9).

Staying clear from the realm of science fiction, Amartya Sen makes the point in the following manner:

> A person who is ill-fed, undernourished, unsheltered and ill can still be high up in the scale of happiness or desire-fulfilment if he or she has learned to have 'realistic' desires and to take pleasure in small mercies.... Considerations of 'feasibility' and of 'practical possibility' enter into what we dare to desire and what we are pained not to get. Our mental reactions to what we actually get and what we can sensibly expect to get may frequently involve compromises with a harsh reality. The destitute thrown into beggary, the vulnerable landless labourer precariously surviving at the edge of subsistence, the overworked domestic servant working round the clock, the subdued and subjugated housewife reconciled to her role and fate, all tend to come to terms with their respective predicaments. The deprivations are suppressed and muffled in the scale of utilities (reflected by desire-fulfilment and happiness) by the necessity of endurance in uneventful survival (Sen 1987, 14-15).

If the considerations brought up by Griffin, Nozick and Sen are correct, it appears that the preference hedonist account of well-being are not satisfactory as accounts of well-being.

It should be pointed out that nothing that has been said so far implies that subjective measures are necessarily inadequate as measures of well-being properly understood. Even if the accounts of well-being presupposed by the psychologists turn out to be false, it is possible that these measures – to some degree, perhaps by chance – represent some other account of well-being, and that this other account happens to be correct. If this is so, however, these measures can hardly be described as "direct" in any interesting sense of the term (like many of the psychologists do). Thus, while it remains possible that subjective measures may be defensible as imperfect measures of well-being, the cannot be defended as more "direct" than traditional economic measures.

The idea that subjective measures can serve as indirect measures of well-being is not completely implausible. Proponents of subjective measures often argue that happiness and other positive mental states are correlated with other characteristics arguably associated with well-being properly understood. Thus, David G. Myers writes:

> Happy people ... are strikingly energetic, decisive, flexible, creative, and sociable. Compared to unhappy people, they are more trusting, more loving, more responsive.... Happy people tolerate more frustration. They are less likely to be abusive and are more lenient. Whether temporarily or enduringly happy, they are more loving and forgiving and less likely to exaggerate or overinterpret slight criticism. They choose long-term rewards over immediate small pleasures.... Moreover, in experiment after experiment, happy people are more willing to help those in need.... Evidence also accumulates that, as the writer of the Proverbs said, " a cheerful heart is a good medicine, but a downcast spirit dries up the bones".... Our body's immune system fights disease more effectively when we are happy rather than depressed (Myers 1992, 20-21).

In sum, Myers concludes: "So, human happiness is both an end – better to live fulfilled, with joy – and a means to a more caring and healthy society" (Myers 1992, 21). Yet, even if higher scores on happiness scales correlate with well-being (properly understood), this does not change the fact that subjective measures cannot be defended as measures of well-being (properly understood) on the grounds that they are more direct than other, e.g. economic, measures.

In spite of the perhaps disappointing conclusions from the point of view of subjective measures, nothing that has been said here should be taken to imply that subjective measures must be inferior to traditional economic measures. As it happens, preference satisfaction accounts according to which what matters are actual preferences are implausible as well. In fact many of the argument against mental state accounts of well-being (such as that of Sen 1987, 14-15, quoted earlier in this section) can also be turned against actual preference satisfaction accounts. While preference satisfaction accounts according to what matters are the preferences that the individual would have under some fairly stringent counterfactual conditions may be more plausible as accounts of well-being, there is little to no evidence that e.g. real income is a good indicator of the preferences real people would have under counterfactual conditions. Not that it is easy to identify the preferences an individual would have under such conditions, but in the presence of ignorance and irrationality there

is little reason to think that people with more money in general succeed in satisfying their ideal preferences to a higher degree than others.

3.8 CONCLUSION

In this chapter, I have argued for a series of connected theses. (a) Accounts of well-being differ radically both across and within disciplines. (b) Proponents of subjective measures have failed to properly acknowledge the fact that accounts differ. (c) The lack of acknowledgement has had a series of unfortunate consequences, including that of misrepresenting the economists' account of well-being (and perhaps of rationality). (d) The plausibility of psychologists' arguments in favor of subjective measures depends on the plausibility of the account of well-being that they use. (e) That account is implausible as an account of well-being. (f) Thus, subjective measures cannot be defended on the grounds that they are more direct than e.g. economic measures. (g) Yet, the discussion does not imply that subjective measures are any less plausible than economic accounts, since preference satisfaction accounts are associated with problems similar to those affecting mental state accounts.

One of the upshots is that questions relating to the proper measure of well-being to be used for purposes of public policy will depend on some fundamentally philosophical questions, most prominently about the nature of well-being. The argument above suggests that the plausibility of both subjective and economic measures will depend, to some extent, on the question of the correct account of well-being. This, in turn, suggests that questions about the correct account of well-being cannot be ignored. This contention will be strengthened in the following chapter, in which I argue that subjective measures cannot be dismissed on measurement theoretic considerations alone.

4.0 THE MEASUREMENT THEORETIC ARGUMENT

Abstract. In this chapter I examine what may be the most common argument against subjective measures of well-being. The argument relies on the claim that the degree to which people are happy or satisfied cannot be measured. As an argument against subjective measures, this one is particularly hard to assess, since most of its central assumptions remain suppressed. Thus, the aim of this chapter is twofold: first, to identify the central assumptions on which the argument rests; second, to assess how convincing it is. I argue that the argument is best understood as based on the following premisses: (a) measurement requires the existence of an observable ordering; (b) actual choices of economic agents constitute such an ordering; and (c) no analogous observable ordering exists in the case of happiness or satisfaction. If this is a correct analysis of the argument, it relies on an empirical assumption – viz. that agents' choices satisfy the axioms of rational choice theory – which is increasingly difficult to defend in light of recent empirical developments in behavioral economics. As a result, I claim, the argument as it stands is unconvincing. Nevertheless, many serious questions regarding the subjective measures remain.

4.1 INTRODUCTION

In this chapter, I examine what may be the most common argument against psychological measures. According to this argument, subjective measures of well-being are inadequate because the degree to which people are happy or satisfied cannot be adequately measured. The implied contrast, of course, is with degrees of preference satisfaction, which according to this line of thought can be measured. The critic recognizes that economists and psychologists operate with different accounts of well-being: whereas economists adopt some kind of preference satisfaction account, psychologists rely on

some mental state account. The idea is that the former permits the development of adequate measures of well-being, whereas the latter do not.

This line of argument – to which I will refer as the *measurement theoretic* argument – appears in Wilfred Beckerman's book *Two Cheers for the Affluent Society: A spirited defense of economic growth* (1974), in which the author defends the use of Gross National Product (GNP) as a measure of welfare. Beckerman recognizes that there are different views of what constitutes welfare, and he acknowledges that some people think welfare is a matter of happiness. Nevertheless, he insists on using GNP as a measure of welfare. Beckerman's main reason is that "[the] concept of happiness is one for which there can be no scientific objective measure" (Beckerman 1974, 53). The passage is clearly intended to suggest an argument against measures of well-being-as-happiness, on the grounds that degrees of happiness do not permit the development of "scientific objective" measures. Nevertheless, the fact that the argument is so common, and that it often presented in a way that suggests that it is of the knock-down variety, makes it well worth our attention.

In this chapter, my primary goal is to assess the plausibility of the argument against subjective measures. Unfortunately, the argument is difficult to assess, because most of its central assumptions remain suppressed; in particular, it is unclear what considerations are supposed to support the contention that happiness (for example) cannot be measured. The first order of business, therefore, is to clarify what the central assumptions of the argument are. In my view, the most convincing way to identify the implicit and unargued assumptions on which economists rely is to situate them in the intellectual tradition of which they are part. Thus, I examine the historical episode during which economics was purged of all references to hedonic psychology. I claim that this episode, which took place around the 1930's, helped shape economists' views about the role of mental states in social science. Next, I examine two research programs that grew out of this period, both of which have had a remarkable influence on the development of modern economics, and in

particular on economists' efforts to measure welfare. I will argue that this historical background sheds a great deal of light on current approaches to welfare measurement.

The measurement theoretic argument is interesting for several reasons. First of all, it represents an obvious challenge for proponents of subjective measures. If the argument is sound and mental states do not permit the development of adequate measures, it would seem that the whole "science of happiness" would suffer a devastating blow. That is, it would seem that subjective measures could not legitimately be used to identify the determinants and distribution of well-being (as the psychologists understand it). For the same reason, it would appear, they could not be used as a guide in the articulation and evaluation of public policy. Thus, the reorientation of public policy envisioned by psychologists working on subjective measures may have to be aborted.

Moreover, it is sometimes argued that accounts of well-being can be judged in part on the grounds of whether they permit the development of adequate measures of well-being. This idea is explicit in a number of prominent contemporary philosophers. James Griffin argues that we cannot "first fix on the best account of 'well-being' and independently ask about its measurement. One proper ground for choosing between conceptions of well-being would be that one lends itself to the deliberation that we must do and another does not" (Griffin 1986, 1). The idea, presumably, is that the measurement of well-being is necessary for the deliberation that we have to do. Similarly, Christine M. Korsgaard maintains that an account of the quality of life may be assessed "for its utility in determining actual political and economic policy – that is, whether it provides accurate enough measures to assess the effects of policy" (Korsgaard 1993, 54). If Korsgaard is correct, the question of what can be adequately measured is eminently relevant to the adequacy of accounts of well-being, and to the plausibility of the ethical theories in which that concept plays a role.

The main conclusion of this chapter is that the measurement theoretic argument remains unconvincing. This is so in part because it relies critically on an empirical premiss – viz. that people's

choices satisfy the axioms of rational choice theory – which is increasingly difficult to maintain in light of recent advances in behavioral economics. The conclusion does not mean, however, that e.g. degrees of happiness in fact are easier to measure than degrees of preference satisfaction. There are, in fact, serious issues associated with subjective measures, but they are quite different from the one suggested by the measurement theoretic argument.

4.2 ECONOMICS AND HEDONIC STATES

In this section I discuss the historical background to economists' skepticism regarding mental states and their measurement. This history is interesting in part because hedonic psychology used to be seen as the very foundation of economic analysis. During the first half of the twentieth century, however, economics was gradually purged of references to hedonic psychology as well as appeals to introspection. Instead, the concept of preference took over the role previously played by concepts like happiness. This historical background will allow us to address include why economists are so skeptical of references to happiness, and what the answer to that question tells us about their attitudes to measurement.

4.2.1 Hedonic psychology as a foundation for economics

In its early development, economic theory was closely tied up with hedonic psychology.[55] During this period, economic theory was based on an account of individual behavior according to which individuals seek to maximize utility, where utility was taken to be some mental state like "happiness" or "pleasure." In this they followed Jeremy Bentham, who wrote: "My notion of man is, that,

[55] Parts of the history I cover here have been discussed by Alexander Rosenberg (1981) and Don Ross (1999).

successfully or unsuccessfully, he aims at happiness, and so will continue to aim as long as he continues to be man, in every thing he does" (quoted in Ross 1999, 34). Similarly for William Stanley Jevons, who claimed:

> Pleasure and pain are undoubtedly the ultimate objects of the Calculus of Economics. To satisfy our wants to the utmost with the least effort – to procure the greatest amount of what is desirable at the expense of the least that is undesirable – in other words, to *maximize pleasure*, is the problem of Economics (in Ross 1999, 89, italics in original).

The focus on conscious experience was, of course, shared by welfare economists. As A. C. Pigou put it in *The Economics of Welfare* (1952 [1920]), "the elements of welfare are states of consciousness and, perhaps, their relations" (Pigou 1952, 10).

Since utility was taken to be a matter of conscious experience, the foundations of economics were often defended on the basis of their introspective "self-evidence." For example, John E. Cairnes (1888) wrote:

> The economist may ... be considered at the outset of his researches as already in possession of those ultimate principles governing the phenomena which form the subject of his study, the discovery of which in the case of physical investigation constitutes for the inquirer his most arduous task (Cairnes 1888, 89-90)

Similarly, "*The economist starts with a knowledge of ultimate causes*. He is already, at the outset of his enterprise, in the position which the physicist only attains after ages of laborious research" (Cairnes 1888, 87, italics in original). Cairnes continues:

> For the discovery of such premises no elaborate process of induction is needed.... It is not necessary to do this—to resort to this circuitous process—for this reason, that we have, or may have if we choose to turn our attention to the subject, direct knowledge of these causes in our consciousness of what passes in our own minds (Cairnes 1888, 88).[56]

[56] See Robbins (1932, 84-85) and Hutchison (1938, 131-133) for more about the role of hedonic psychology in early modern economics.

In Cairnes' view, quite clearly, the truth of fundamental principles of economics is immediately obvious.

4.2.2 The purging of hedonic psychology from economics

During the interwar era, however, an increasing number of economists objected to the notion that economics should make reference to conscious states. Those same economists also rejected the idea that introspection was a scientifically acceptable means to explore such states. Many of these authors were quite clearly inspired by the methodological strictures of logical positivism in philosophy, behaviorism in psychology, and operationalism in physics. Moreover, at least some economists appear to have grown disappointed with the meager results (in terms of theories with predictive power) of accepted economic practice. T. W. Hutchison (1938) gives voice to this sentiment when he writes:

> It is possibly very encouraging for the economist to hear that compared with the natural scientist the psychological method saves him "ages of laborious research," but it is curious and a pity that this huge start has not enabled him to formulate any considerable body of reliable prognoses such as the natural sciences have managed to achieve (Hutchison 1938, 132).

Thus, great a number of economists set out to put their discipline on firmer methodological ground, and at the same time to improve the predictive power of their theories. They did so by rejecting hedonic psychology and instead emphasizing that which can be "objectively observed." As we will see below, the ambition has influenced the development of modern economics – and especially the measurement of welfare – in profound ways.

One of the best representatives of historical episode in which economics was purged of hedonic psychology is Lionel Robbins, author of the spectacularly influential *An Essay on the Nature and Significance of Economic Science* (1984 [1932]). This book has a two-fold goal:

> In the first place, it seeks to arrive at precise notions concerning the subject matter of Economic Science and the nature of the generalisations of which Economic Science consists. Secondly it attempts to explain the limitations and the significance of these generalisations, both as a guide to the interpretation of reality and as a basis for political practice (Robbins 1984, xli).

Here, I will focus on Robbins' views about the nature of economic generalizations, especially those relating to individual behavior (Robbins 1984, chapter IV). The views of Robbins are worth exploring because of their extraordinary influence on contemporary economics.

Robbins explicitly rejects the idea that economic theory in any way depends on the truth of particular psychological doctrines (Robbins 1984, 83). He writes: "The borderlines of Economics are the happy hunting-grounds of minds averse to the effort of exact thought, and, in these ambiguous regions, in recent years, endless time has been devoted to attacks on the alleged psychological assumptions of Economic Science" (Robbins 1984, 83). He acknowledges that past economists have invited the interpretation that economics is based on hedonic psychology. Nevertheless, he insists that recent theory has no "essential connection with psychological hedonism, or for that matter with any other brand of *Fach-Psychologie*" (Robbins 1984, 85). Indeed, the theory "is capable of being set out and defended in absolutely non-hedonistic terms" (Robbins 1984, 85).

In Robbins' view, economics is based instead on the notion of *relative valuation*. As he writes:

> As we have seen already, all that is assumed in the idea of the scales of valuation is that different goods have different uses and that these different uses have different significances for action such that in a given situation one use will be preferred before another and one good before another (Robbins 1984, 85-86).

Notice how Robbins slides from talk about relative valuation to talk about preference. It is not clear if Robbins identifies the two, but we can be certain that they are closely connected. Robbins does not assume that it is impossible to account for relative valuations (or preferences), but he insists that this enterprise is no part of economics. He writes: "Why the human animal attaches particular values

in this sense to particular things, is a question which we do not discuss. That is quite properly a question for psychologists or perhaps even physiologists" (Robbins 1984, 86).

Robbins acknowledges that his interpretation assumes that people are rational in a certain sense (Robbins 1984, 90-91). This does not contradict the general argument, however, since in Robbins' view rationality can be defined by reference to an agent's choices only. He writes that

> ... in so far as rational is taken to mean merely "consistent," then it is true that an assumption of this sort does enter into certain analytical constructions. The celebrated generalisation that in a state of equilibrium the relative significance of divisible commodities is equal to their price, does involve the assumption that each final choice is consistent with every other, in the sense that if I prefer A to B and B to C, I also prefer A to C (Robbins 1984, 92).

Consistency, in Robbins' terms, is the property typically referred to as transitivity. Note how easily Robbins goes from talk about preferences to talk about choice. While Robbins makes it clear that he does not identify the two (Robbins 1984, 87-88), he obviously assumes that they are closely connected. Presumably, preferences – at least to a great extent – determine choices, so that if preferences are consistent so are choices and *vice versa*.

In summary, Robbins believes that modern economics – in spite of its historical origins – is entirely independent of hedonic psychology. Instead, it is based on the notion of relative valuation, a concept that is closely connected with (if not identical to) that of preference. Moreover, there is some causal connection between preferences and choices, such that choices are consistent just in case preferences are consistent, and choices in some sense mirror preferences. It follows that we can learn about people's preferences by studying their choices. As Robbins points out, the argument assumes that relative valuations, preferences, and choices are all consistent in a certain sense, viz. that they are transitive.

Since the publication of *The Nature and Significance of Economic Science*, a number of economists have aspired to provide methodological underpinnings for talk about preference (and utility) along

the lines suggested by Robbins. No doubt, typically the goal was both to render economic method more in line with contemporary methodological strictures, and to improve the predictive accuracy of the theory. In this tradition, some people aspired to offer a complete account of "preference" – as in a definition of the term – such that a particular method can be justified. In other cases, people were happy to offer an account of a method whereby preferences can be identified, while remaining ontologically non-committal, as it were. In the following two sections, I will describe two such attempts, one of each kind, both of which have been widely influential. As we will see, in spite of their differences, these accounts share a number of features. Both imply that preferences can be unambiguously identified by studying agents' choices. Moreover, both require that these choices satisfy some consistency condition.

4.2.3 Revealed preference theory

Revealed preference theory was developed by Paul Samuelson in a series of publications in the thirties and forties (1938a, 1938b, 1947, 1948). Though it first appeared when Samuelson was only 23 years old, the theory was one of the achievements that would earn him the 1970 Nobel Memorial Prize in Economics. In Samuelson's view, the scientific respectability of utility theory was drastically improved when it was shown how every statement about utilities could be rewritten as a statement about preferences. In this view, utility is defined as an index of preference satisfaction, so that a person is taken to assign more utility to alternative *A* than to alternative *B* if and only if she prefers *A* to *B*. Samuelson considers the proof a step forward for utility theory, consumer choice theory, and neo-classical economics more generally, since he judges statements about independently existing utilities scientifically illegitimate. However, he reports a certain suspicion concerning the notion of

preference itself. In fact, Samuelson fears that it may be no more legitimate than the notion it was meant to replace.[57]

In defense of utility theory, Samuelson attempts to show how every statement about preferences can be rewritten as a statement about observable choices. The tacit assumption, of course, is that statements about choices are legitimate as they stand. Formally speaking, Samuelson proves that is possible to construct an ordering which satisfies the axioms of rational preference on the basis of an agent's choices, provided those choices satisfy certain consistency conditions (cf. Hausman, 2000, 100-101). The exact formulation of the consistency conditions varies, but can be phrased as follows: "if an individual selects batch one over batch two, he does not at the same time select two over one" (1938a, 65).[58] In 1948, Samuelson summarizes his contribution as follows: "A decade ago I suggested that the economic theory of consumer's behaviour can be largely built up on the notion of 'revealed preference' ... the individual guinea-pig, by his market behaviour, reveals his preference pattern—if there is such a consistent pattern" (1948, 243). Samuelson concludes: "The whole theory of consumer behaviour can thus be based upon operationally meaningful foundations in terms of revealed preference" (1948, 251).

The literature on revealed preference presents two main interpretations of Samuelson's theory.[59] According to the first interpretation, the theory provides a novel, operational definition of the notion of preference by identifying it with the notion of choice. On this view, "Person *P* prefers *A* over *B"* is *defined* as "Person *P* chooses *A* over *B* when both alternatives are available." Accordingly, a preference for red wine over white does not *cause* you to choose red rather than white at dinner parties. Instead, to say that you prefer red to white just *means* that you choose red rather

[57] This paragraph draws on Samuelson (1938a, 61-62) and (1947, Chapter V).

[58] The textbook version of the assumption is often referred to as the Weak Axiom of Revealed Preference (WARP).

[59] See Sen (1982, 55-57) and Hausman (2000, 100-101).

than white when both are available. According to the second interpretation, in contrast, Samuelson's theory does not reject preferences as subjective mental states. Rather, the theory shows how an observer can legitimately use a person's choices to make inferences concerning her subjective preferences. On this view, the connection between choice and preference is synthetic, rather than analytic.

Here, I discuss revealed preference theory under the first, operationalist interpretation. I do so for several reasons. First, the operationalist interpretation is most likely to be Samuelson's own. There are strong grounds for believing that operationalism in the philosophy of science motivated his original work, and that he drew on Percy W. Bridgman's operationalist manifesto *The Logic of Modern Physics* (1927), which appeared some ten years before the 1938 paper. As Hausman says, "Paul Samuelson sketched out an 'operationalist' program for economic theory that apparently offered a new empirically respectable way of doing economics" (1992, 156). Thus, the operationalist interpretation of revealed preference appears to be the original one.[60] Moreover, the interpretation is common in literature on both economics, philosophy of science, and ethics. Ian M. D. Little, for instance, defends Samuelson's analysis as "scientifically more respectable," since "[if] an individual's behaviour is consistent, then it must be possible to explain that behaviour without reference to anything other than behaviour" (1949, 97). John Hicks asserts that his "econometric theory of demand does study human beings, but only as entities having certain patterns of market behavior; it makes no claim, no pretense, to be able to see inside their heads" (1956, 6). Robert Sugden writes that the correspondence between preference and choice according to revealed preference theory "is a *definition* of preference, not just an assumption about preference" (Sugden, 1992, 32). In his philosophical critique of the theory, Alexander Rosenberg clearly targets the operationalist interpretation:

[60] Cf. Samuelson's ambition to provide "operationally meaningful foundations" (Samuelson 1948, 251).

> Revealed preference theory tells us that ... we should interpret ordinal utility theory as a behaviorist would, instead of treating preference as causing behavior and as 'revealed' by it. We should *redefine* '*X* prefers *A* to *B*' to *mean* '*X* actually chooses *A* when both *A* and *B* are available' (1992, 119-120, my italics).

Finally, Ken Binmore asserts: "The theory of revealed preference ... *deduces* a consistent person's preferences from the decision he makes ... At the analytical level, it is *tautological* that *homo economicus* maximizes all the time" (1994, 27, italics in original). In sum, the operationalist interpretation of revealed preference theory deserves to be taken seriously, both because it appears to be the original one, and because it is common in the literature.

What exactly characterizes an operational definition? Consider a paradigmatic example of operational definition, such as that of 'length':

> *X* is one meter long if and only if, when a meter rod is aligned with *X*, the edges of *X* coincide with the ends of the rod. (Def. of 'length')[61]

The definition specifies an operation, by giving the conditions under which an operation can be said to be performed, and a characteristic response. The term defined – the definiendum – is said to apply just in case the characteristic response occurs under the specified test conditions. It is understood that the test conditions and the characteristic response should be observable. That is, under the right conditions it should be a purely observational matter to determine whether the conditions obtain and whether the response occurs.

Analogously, we can define the notion of preference in the following manner:

> *P* prefers *A* to *B* if and only if, when both *A* and *B* are available to *P*, *P* chooses *A* rather than *B*. (Def. of 'preference')

[61] Operational definitions of 'length' are discussed e.g. in Bridgman (1927, 9 ff.). See Hempel (1965 [1954], 123-128) for a clear explication of Bridgman's program.

Again, the definition specifies an operation and a response, and the definiendum is said to apply just in case the response occurs when the conditions obtain. Again, it should be a purely observational matter to determine whether the conditions obtain, and whether the response occurs.

Revealed preference theory, as characterized here, has appealed to economists for a number of different reasons. First, it provided an answer to important questions concerning the ontological status of preferences. If we accept revealed preference theory, we need not be committed to preferences "inside the head," any more than we are committed to independently existing utilities. The revealed preference theorist wishes to be at least agnostic about such entities, and want to commit himself only to the existence of choice behavior. Second, the theory allowed economists to answer questions about the connection between preferences and choice behavior. On this view, the connection is logical, rather than e.g. causal, since claims about preferences are deductively entailed by claims about choices. Third, revealed preference theory was taken to validate a particularly simple methodology for empirical economic research. When the revealed preference theorist aspires to chart people's preferences over some set of alternatives, she will simply study subjects' choices between the alternatives and directly deduce the preferences. No auxiliary theories or complex inferences are required.

In sum, Samuelson's contribution can be characterized as a radical effort to excise every remnant of hedonic psychology from economic theory. In Samuelson's view, there is nothing scientifically illegitimate about talk of preference and utility *per se*. However, all such talk must ultimately be reducible to talk about choice, where choice is understood as overt, directly observable behavior (rather than, say, the making up of one's mind). Samuelson explicitly recognizes that his account is based on the assumption that an agent's choices are consistent, in a sense that can be spelled out in a formally precise way. Note the central importance of the consistency assumption in

Samuelson's proposal. In fact, if an agent's choices violate the consistency assumption, Samuelson's account gives us no grounds to attribute preferences to her at all.

4.2.4 Measurement theory as the foundation for preference

While revealed preference theory has been falling out of fashion in recent years, a more enduring attempt to provide solid methodological foundations for talk about utility and preference is based on the theory of measurement. Conceptually, of course, in this view the measurement of utility is a specific instance of measurement *tout court*. Historically, however, the two emerged in conjunction. Indeed, the development of the theory of measurement was to a significant degree influenced by problems associated with the measurement of utility (see Krantz et al. 1971, 9). The theory behind the measurement of utility was developed by John von Neumann and Oskar Morgenstern (1944), Leonard J. Savage (1954) and others,[62] drawing on Frank P. Ramsey (1931). The theory of measurement, in its modern shape, was first articulated by Dana Scott and Patrick Suppes (1958) but received what may be its canonical statement in the three-volume *Foundations of Measurement,* the first volume of which appeared as David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky (1971).[63] Here, I also rely on the retrospective article 'From Indices to Mappings: The representational approach to measurement' by Krantz (1991).

Krantz (1991) notes that there are, broadly speaking, two approaches to measurement in the behavioral and social sciences. In his words: "One, which may be termed the psychometric approach, introduces latent [unobservable] variables to explain behavioral orderings. The second ... treats the numerical representation of behavioral orderings axiomatically" (Krantz 1991, 2). For the

[62] Cf. Krantz et al. (1971, 9)

[63] Subsequent volumes was published as Suppes, Krantz, Luce and Tversky (1989) and Luce, Krantz, Suppes, and Tversky (1990).

obvious reason, I will refer to the second approach as the *behavioral-representational*. The focus of Krantz and co-authors is on the latter, the theory of which is often referred to simply as the theory of measurement. In their introductory chapter, Krantz et al. (1971) offer a concise statement of the purpose and nature of measurement as they see it. They write:

> When measuring some attribute of a class of objects or events, we associate numbers (or other familiar mathematical entities, such as vectors) with the objects in such a way that the properties of the attribute are faithfully represented as numerical properties (Krantz et al. 1971, 1).

Thus, at one level of abstraction, measurement is the process of attributing numbers to object so as to represent some relevant property of the objects at hand. More specifically, and in more technical language:

> From this standpoint, measurement may be regarded as the construction of homomorphisms (scales) from empirical relational structures of interest into numerical relational structures that are useful. Foundational analysis consists, in part, of clarifying (in the sense of axiomatizing) assumptions of such constructions (Krantz et al. 1971, 9).

In what follows, I will try to explain what this means, and how it applies to the measurement of utility.

It helps to think of measurement in the context of an actual example, so let us follow Krantz et al. (1971) and consider (again) the case of length measurement. In their words:

> Suppose that we have a set of straight, rigid rods whose lengths are to be measured. If we place the rods a and b side by side and adjust them so that one is entirely beside the other and they coincide at one end, then either a extends beyond b at the other end, or b beyond a, or they appear to coincide at that end also. We say, respectively, that a is longer than b, b is longer than a, or that a and b are equivalent in length. For brevity, we write, respectively, $a \succ b$, or $b \succ a$, or $a \sim b$. Two or more rods can be *concatenated* by laying them end to end in a straight line, and so we can compare the qualitative length of one set of concatenated rods with that of another by placing them side by side, just as with single rods. The concatenation of a and b is denoted $a \circ b$ and the observation that c is longer than $a \circ b$ is

> denoted $c \succ a \circ b$, etc. Many empirical properties of length comparison and of concatenation of rods can be formulated and listed, e.g., $\succ$ is transitive; $\circ$ is associative; if $a \succ b$, then $a \circ c \succ b$; etc. (Krantz et al. 1971, 2).

The basic idea is the following. A set A of objects, in this case a set of rods, can be ordered e.g. with respect to length. We can figure out how various rods are related to each other with respect to length by applying the simple operation described above. This ordering of rods will as a matter of fact satisfy a number of conditions, like transitivity. These conditions can be identified by empirical study, and, incidentally, are most conveniently expressed as set of axioms (cf. Krantz et al. 1971, 6). The axioms can be seen as a set of empirical laws (Krantz et al. 1971, 13). Thus, "fundamental measurements are based on certain qualitative physical laws" (Krantz 1972, 1428).

The example illustrates what Krantz et al. mean by an *empirical relational structure*.[64] A *relational structure* is a set of objects along with relations on that set (cf. Krantz et al. 1971, 8). In this case, we have a set (which we will call A) of rods. Moreover, we have two relations of interest: $\succ$, which is a binary relation, and $\circ$, which is ternary, holding between a, b, and $c = a \circ b$ (Krantz et al. 1971, 8). Such an empirical relational structure is referred to as $\langle A, \succ, \circ \rangle$. An empirical relational structure contrasts with a *numerical relational structure*, which is a set of mathematical objects like numbers along with relations on that set.

Given an empirical relational structure, we want to assign numbers $\phi(a)$, $\phi(b)$, etc. to rods a, b, etc. in such a way that the following two conditions are satisfied (see Krantz et al. 1971, 5). First, we require that the number assigned to a be greater than the number assigned to b just in case a in fact is longer than b. That is, $\phi(a) > \phi(b)$ if and only if $a \succ b$. Second, we require that the numbers assigned be additive with respect to concatenation. That is, $\phi(b \circ c) = \phi(b) + \phi(c)$.

[64] Krantz (1972) refers to them as *measurement structures*, and Krantz (1991) as *qualitative structures*.

If we succeed in assigning numbers such that these two conditions are satisfied, the function $\phi(\cdot)$ is a *homomorphism* from the empirical relational structure into a numerical relational structure $\langle \Re, >, + \rangle$. This means that $\phi(\cdot)$ takes elements of A into the set $\Re$ of real numbers in such a way that the corresponding relationships are preserved. That is, $a \succ b$ if and only if $\phi(a) > \phi(b)$, and $c = a \circ b$ if and only if $\phi(c) = \phi(a) + \phi(b)$. If a homomorphism $\phi(\cdot)$ is one-one, we say that it is an *isomorphism*. This will occur if no two rods are equal with respect to length, so that each rod gets assigned a unique number. Especially in the empirical literature, homomorphisms are often referred to as *scales*. This explains the view of Krantz et al. (1971, 9) that the process of measurement can be seen as the process of constructing homomorphisms, or scales, from empirical to numerical relational structures.

Not every empirical relational structure will allow the construction of a homomorphism, however. As a result, we will want to know what conditions must be satisfied by the empirical relations $\succ$ and $\circ$ on A such that a function $\phi(\cdot)$ with desired properties can be constructed (cf. Krantz et al. 1971, 6). As Krantz et al. write: "A measurement procedure certainly is not adequately understood if it depends on properties that are not explicitly recognized" (Krantz et al. 1971, 6). At this point, it is possible to explore formally what axioms are necessary and sufficient for it to be possible to construct a function $\phi(\cdot)$ that satisfies the properties identified above (Krantz 1971, 8). The answer to a question such as this one is given by a representation theorem, which "asserts that if a given relational structure satisfies certain axioms, then a homomorphism into a certain numerical relational structure can be constructed" (Krantz et al. 1971, 9).

The measurement-theoretic point of view aspires to be – in a certain sense – ontologically non-committal. When measurement theorists talk about length, hunger, frustration, risk aversion, and so on, Krantz notes, it may seem as if they "introduce ontological presuppositions" and

presuppose the existence of such things (Krantz 1991, 3). However, he argues, in actual practice measurement theorists do not need to take a position on the issue of the ontological status of such entities, since everything they need to assume is the existence of an empirical ordering that satisfies certain conditions. As Krantz puts it, "in most cases it seems that one expects at least a useful sort of ordering of objects or situations or organisms or social entities" (Krantz 1991, 3). In order to justify talk about length, on this view, everything we need to assume is that there exists an empirical structure of a set of rods and relations that satisfy certain conditions. It is important to notice, by the way, that measurement theory is not an attempt to provide operational definitions of theoretical terms. Krantz et al. write that to treat "indirect measures" as "objective definitions of unanalyzed concepts" is a temptation that has to be resisted (1971, 32).

The theory of measurement may seem entirely too abstract to shed any light on real life examples such as the measurement of utility (cf. Krantz et al. 1971, 9). However, as we will see, according to measurement theorists the measurement of utility is a straightforward application of the account described above. In brief, the assumption underlying the measurement of utility is that a choice structure – a set X of options and a choice relation R on X – is an empirical relational structure satisfying certain axioms. Though there are different ways to approach the topic, typically X is the set of all possible acts or bundles, and R is a binary relation such that aRb means that a is chosen over b in a pair-wise choice.

Krantz (1991, section 3) discusses the example of utility quite explicitly. He writes:

> Since 1960, there seems to have been general agreement concerning two main points about the measurement of utility. First, the empirical ordering underlying utility is determined by actual choices; that is, the choice of one act over others is represented by a utility assigned to the chosen act that is higher than the utilities assigned to the other acts (Krantz 1991, 28).

The second point has to do with whether utility scales are ordinal or cardinal, which is a topic I will avoid going into here. Krantz goes on: "The first of these points reflects the view that it is actual choices that are the most trustworthy and most important data of a behavioral science" (Krantz 1991, 28). Typically, of course, there is an assumption that choices reflect preferences over the various options. Anyway, as Krantz notes, "most utility theories cling to the idea that the ordering is based on observation of choice behavior" (Krantz 1991, 28-29).

In order for the representation theorem to work, the measurement theorist assumes that the choice structure $\langle X, R \rangle$ satisfies some set of conditions. Either one of several different sets of axioms will do the trick. However, there are some conditions that are shared by all axiomatizations of utility. As Krantz et al. (1971, 21-22) point out, transitivity is a *necessary* condition, in the sense that it is mathematically necessary for the representation theorem to work. This claim is proved by the following simple argument. If aRb and bRc, the fact that $\phi(\cdot)$ is a homomorphism implies that $\phi(a) > \phi(b)$ and $\phi(b) > \phi(c)$. Thus, $\phi(a) > \phi(c)$, and since $\phi(\cdot)$ is a homomorphism, aRc. QED.

Just like in the case of the measurement of length, the axioms are seen as empirical (descriptive) laws, in this case, laws of choice. Thus, for instance, the axioms articulated by von Neumann and Morgenstern (1944) "constitute a set of qualitative laws for 'rational' decisions among risky options" (Krantz 1972, 1428). Very often, of course, these axioms are treated as normative laws of rational choice. In the present context, however, the normative status of the axioms is irrelevant. The point here is that the representation theorem requires that they be true descriptive laws. That is, for the theorem to apply the choice structure must *in fact* satisfy the axioms, viz. that the axioms must be *true* of the empirical relational choice structure.

Although the theory of measurement is a complicated affair, and though the presentation in this subsection is brief, we can already see how the measurement of utility is supposed to relate e.g. to the measurement of length. Instead of a set of rods we have a set of bundles or acts. Instead of an

ordering determined by comparisons of rods placed side-by-side, we have an ordering determined by the choices of some agent. In order to allow the construction of a representation theorem, we need to identify a set of axioms, which can be seen as empirical, descriptive laws of choice. One of these laws is transitivity, that is, if *a* is longer than *b* and *b* is longer than *c*, then *a* is longer than *c*. Just like in the case of length, measurement theory is supposed to allow us to remain agnostic about the existence of utilities (and preferences, presumably). Thus, the measurement theorist claims to provide methodological foundations for talk about utility and preference without "pretending to look into the head of the agent." All the measurement theorist assumes is that choices determine an empirical relational structure with the appropriate properties, that is, that actual choices satisfy the appropriate axioms. Note, again, the importance of the consistency condition. If this condition is not in fact satisfied, measurement theory gives us no grounds whatsoever for constructing a utility function on the set of available options.

4.2.5 Discussion

In this section, we have seen that economists' rejection of hedonic psychology was based on two separate ideas. Firstly, there was the idea that references to mental states are scientifically illegitimate and constitute an obstacle to scientific progress. Secondly, there was the idea that references to mental states are unnecessary, since the concept of preference (or relative evaluation) is sufficient to provide the foundations for economic theory. Both ideas were forcefully defended during the 1930's, and have remained remarkably influential. In particular, as we have seen, these ideas motivated the more recent development of revealed preference theory and the theory of measurement. Both these research programs were designed, in part, to provide methodological foundations for talk about welfare. Consequently, this history can be expected to shed a great deal of light on current ideas about welfare measurement.

4.3 THE ARGUMENT AND ITS PLAUSIBILITY

4.3.1 The argument explicated

In this section I suggest that contemporary economists' attitudes are usefully seen in light of the historical background discussed in the previous section. More specifically, my claim is that economists – knowingly or not – can be seen as working within the behavioral-representational approach to measurement. This claim is *prima facie* plausible because the theory of measurement, i.e. the behavioral-representational approach, is an important outgrowth of the extraordinarily influential tradition of Robbins and Hutchison, and because the development of axiomatic utility theory – which remains absolutely central to modern economics – was inspired in part by problems of measurement (Krantz et al. 1971, 9). The claim receives additional support, I would argue, by the fact that my interpretation has a great deal of explanatory power (see next subsection).

Assuming that the approach of contemporary economists can be usefully seen in light of the behavioral-representational approach to measurement, it is a straightforward matter to identify the central premisses of the argument against subjective measures. The background reviewed above suggests an argument along the following lines. First:

(1) Measurement requires the existence of an observable ordering.

As we saw above, this is a central assumption of the theory of measurement. Indeed, if the relation of interest fails to satisfy e.g. transitivity, then the theory gives us no grounds for proceeding with assigning numbers.[65] Second:

(2) The (observable) choices of economic agents constitute such an ordering.

[65] The claim that an observable ordering is a necessary condition for measurement is further strengthened by the Krantz passage quoted below (section 4.3.3).

This assumption, in some form or other, was endorsed both by Robbins, Samuelson, and the measurement theorists. In the language of the latter, assumption (2) is equivalent to the following:

(2') The (observable) choices of economic agents satisfy the axioms of rational choice theory.

Thus, assumption (2') says that agents' choices in fact satisfy the "laws" of rational choice. Moreover, an assumption that is shared by all the authors discussed in the previous section is the following:

(3) The (observable) choices of economic agents reflect their preferences, in the sense that *A* is chosen over *B* just in case *A* is preferred over *B*.

This assumption, though implicit, is clearly present both in Robbins, Samuelson, and Krantz et al. Its epistemological status differs, of course. Samuelson takes the assumption to be a definitional truth rather than an empirical one. By contrast:

(4) There is no corresponding ordering in the case of the measurement of happiness, satisfaction, and so on.

Thus, (4) says that there is no observable ordering on which measurement can be based. This claim appears to be completely unargued, but it is required for the argument to be at all effective.

Together, these claims help us understand what I call the measurement theoretic argument, and in particular the proposition that happiness cannot be measured whereas preference satisfaction can. Claims (1) and (4) together imply that degrees of happiness cannot be measured. Meanwhile, claims (1) through (3) imply that degrees of preference satisfaction can.

4.3.2 The explanatory power of my interpretation

In this section, I will argue that my interpretation – including the hypothesis that modern economists can be seen as operating within the behavioral-representational approach to measurement – has great explanatory power. First of all, I believe the hypothesis helps explain the

type of data that economists use in order to construct their measures of welfare. Thus, it explains why economists favor measures that take data about economic transactions – like market choices – as their starting point. The central idea behind most attempts by economists to measure well-being is neatly expressed by Richard E. Just, Darrell L. Hueth, and Andrew Schmitz (1982). They begin by pointing out that utility is not observable, and add:

> In most practical situations the applied welfare economist can, at best, observe income and consumption decisions at various prices and then, on the basis of these economic transactions, try to compute some money-based measure of welfare effects (Just et al. 1982, 69).

Though the economic approaches to the measurement of well-being differ in many ways, almost all of them rely on observable transactions in order to represent utility. In particular, the hypothesis helps account for economists' skepticism regarding questionnaire studies.

This skepticism regarding data about anything other than market-based choice is well documented. In a discussion of dominant attitudes among economists, Amartya Sen comments:

> Choice is seen as solid information, whereas introspection is not open to observation.... Much of economic theory seems to be concerned with strong, silent men who never speak! One has to sneak in behind them to see what they are doing in the market, etc., and deduce from it what they prefer, what makes them better off, what they think is right, and so on (Sen 1982, 9).

Similarly: "Much of the empirical work on preference patterns [and therefore welfare] seems to be based on the conviction that [non-verbal] behaviour is the only source of information on a person's preferences" (Sen 1982, 71).[66] These attitudes can, at least to some extent, be accounted for by

[66] Incidentally, Sen is critical of this attitude. He writes: "That behaviour is a major source of information on preference can hardly be doubted, but the belief that it is the only basis of surmising about people's preferences seems extremely limiting" (Sen 1982, 71). He notes that there are problems associated e.g. with "the questionnaire method," but adds that there are problems with using choice behavior as well (Sen 1982, 71).

pointing out that economists work within the behavioral-representational approach to measurement, which puts a very strong emphasis on observable orderings, like that imposed on a set of alternatives by an agent's choices.

Second, the hypothesis sheds some light on the manner in which economists go about defending their measures. Psychologists are sometimes curious about the fact that economists do not try to establish the validity and reliability of their measures in the manner of the psychometric approach. In the spirit of measurement theory, however, what you need to do to show that a given measure represents utility (or welfare, which comes down to the same thing) is to show that it is a homomorphism, which is a rather different task. Economists tend to assume that the use of a given measure has been justified when it has been shown that it is based on market choices assumed to satisfy the relevant axioms, in conjunction with a formal proof that shows that the measure is an index of preference (i.e. a utility function). Because this procedure, if successful, in fact establishes that the measure is a homomorphism, the behavioral-representational approach helps explain why the procedure is followed.

Third, the hypothesis suggested here helps explain economists' strict adherence to preference satisfaction accounts of well-being. The adoption of a preference satisfaction account of well-being was part of the project of showing that references to mental states were unnecessary, and that the concept of preference would suffice to provide foundations for economic theory. Economists like Robbins, for sure, considered the claim that preference satisfaction can be measured a major consideration in favor of preference satisfaction accounts of well-being. It is plausible to assume that this sentiment remains accurate today, though in most cases it is significantly less explicit than in Griffin and Korsgaard.

Finally, the hypothesis goes a long way toward accounting for economists' skepticism toward subjective measures. The belief that there is no observable ordering that can be used as a basis for

the measurement of happiness, e.g., accounts for the belief that it cannot be measured. I imagine these attitudes go a long way toward accounting for Beckerman's criticism. He and many others, following the behavioral-representational approach, believe that a proper measure is based on a behavioral, observational ordering, and is skeptical of any measure that has not been shown to be a homomorphism.

4.3.3 Is the argument convincing?

Assuming that I have correctly identified the assumptions that go into the measurement theoretic argument against subjective measures, I now proceed to examine its plausibility. I will focus the proposition that mental states like happiness do not permit the development of adequate measures whereas preference satisfaction does. As we saw above, this claim is based on the notion that measurement requires the existence of an observable ordering.

First, note that many economists argue that well-being consists in the satisfaction of those preferences that the agent would have under some specified counterfactual conditions. Yet, there is some serious tension – not to say contradiction – between adopting the behavioral-representational approach to measurement, and adopting a preference-satisfaction account of well-being according to which what matters are ideal preferences. While ideal preference satisfaction accounts (see section 3.2.2) may be more plausible as accounts of well-being, the preference ordering that you would have under some counterfactual conditions – or, the choices that you would make under those conditions – are unobservable by design. Thus, if well-being is understood to be a matter of the satisfaction of ideal preferences, there is no observable ordering that could serve as the basis for measurement.

So let us instead consider the actual preferences that you do have. Do these preferences imply the existence of an observable ordering? In fact, Krantz (1991) discusses this issue.[67] Krantz is impressed by empirical results that seem to suggest that people systematically violate the axioms of rational decision-making. Under the heading "The Myth of Utility" (Krantz 1991, 28), Krantz cites a series of empirical results, and notes: "Choice does indeed depend on the method of testing ... and depends especially on how options are framed" (Krantz 1991, 32). What Krantz' calls the "myth of preference" is, he says, "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, 35). In Krantz' view, agents' choices fail to conform to assumptions of rational choices theory, and therefore these agents do not have preferences; hence the use of the term "myth." This view, if correct, would pose a serious obstacle to the notion that welfare is a matter of the satisfaction of one's preferences.

From our vantage point, if anything, the case for the truth of the axioms of rational choice theory is even weaker than it was 15 years ago, 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. As Matthew Rabin (2002) 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, 662). Similarly, Tversky writes that "if different methods of elicitation give rise to different choices," then "it is difficult to defend the proposition that a person has a well-defined preference order (or equivalently a utility function)" (Tversky 1996, 189).

[67] A similar conclusion has been drawn by Daniel Kahneman, Ilana Ritov, and David Schkade (1999).

So long as we insist on using the behavioral-representational approach to measurement, therefore, it is hard to maintain that degrees of preference satisfaction can be measured. 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 preference satisfaction, 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, then, the central proposition underlying the economic argument against subjective measures appears to be false. If subjective measures fail for the reason identified by Beckerman (1974) and others, economic measures fail too, and for much the same reason.

4.4 THE PSYCHOLOGISTS' DEFENSE OF THEIR MEASURES

If the psychologists do not defend their measures by reference to observable orderings, how do they do it? If they do not adopt the behavioral-representational approach to measurement, what approach do they use? Answers to these questions promises to shed light both on the arguments offered by psychologists in favor of the adequacy of their measures, and on the cogency of economists' criticism of them.

4.4.1 The psychologists' approach to measurement

In this section, I will explore (in slightly greater detail) how the psychologists defend the adequacy of their measures of well-being, and, more generally, what their approach to measurement is. Here, I will rely on the 1999 review article by Diener, Eunkook M. Suh, Richard E. Lucas, and Heidi L. Smith (1999). These authors recognize that there are methodological concerns, among other things,

related to the use of global self-report measures (that is, self-report measures related to life as a whole as opposed to some aspect of it). Diener et al. (1999) respond:

> These measures do possess adequate psychometric properties, exhibiting good internal consistency, moderate stability, and appropriate sensitivity to changing life circumstances. Furthermore, global reports show a moderate level of convergence with daily mood reports, informant reports, spouse reports, and recall for positive versus negative life events. People who score high on global life satisfaction are less likely to attempt suicide and to become depressed in the future (Diener et al. 1999, 277-278).[68]

Thus, in the opinion of Diener et al. (1999), there are several lines of argument that converge on the conclusion that subjective measures (based on self-reports) are adequate. A particularly important reason is that subjective measures are positively correlated with phenomena that we would a priori expect to be positively correlated with happiness (such as spouses' assessments of happiness), and negatively correlated with phenomena that we would a priori expect to be negatively correlated with happiness (such as suicide rates).

This mode of reasoning – along with the reference to "adequate psychometric properties" – strongly suggest that these psychologists operate within the psychometric approach to measurement. As Krantz (1991) puts it (see quote in section 4.2.4 above), this approach "introduces latent [unobservable] variables to explain behavioral orderings" (Krantz 1991, 2). In order to develop a better idea of the nature of the psychometric approach, it is reasonable to have a look at standard textbooks such as the most recent edition of *Psychometric Theory* (Nunnally and Bernstein 1994), or *The New Psychometrics* (Kline 1998).

A central term of the psychometric approach is *reliability*. As Kline puts it, the notion of reliable measurement means that it is "without variation regardless of when the measurement is made or who makes the measurement, provided only that the individual [i.e. the person taking the

[68] References have been omitted.

measurement] is sane, in possession of his or her faculties and trained to use the instrument" (Kline 1991, 26). There are at least two kinds of reliability. A measure has *test-retest reliability* insofar as it yields "the same score for each subject when he or she takes the test on another occasion, given that their status on the variable has not changed" (Kline 1991, 29). A test has *internal consistency* insofar as "each item [of the test administered] measures the same variable" (Kline 1991, 30). A desire to prove the reliability of subjective measures, I would suggest, explains both the concern with stability over time (which is in essence the same thing as test-retest reliability) and internal consistency. If total test scores varied too much over time, or individual items of the test diverged too much from each other, there would be reason to think that the measures were unreliable.

Another central term in the psychometric approach is that of *validity*. As Kline puts it: "A test is said to be valid if it measures what it purports to measure" (Kline 1991, 34; cf. Nunnally and Bernstein 1994, 83). Presumably, the locution "what [the test] is supposed to measure" refers to the attribute that the author or user of the test (as the case may be) takes the test to measure. Again, there are several types of validity, including *face validity, concurrent validity, predictive validity,* and *content validity* (cf. Kline 1991, 34-37) Here I will focus on *construct validity,* which is the form of validity most relevant to the present purposes. This idea, which was introduced by Cronbach and Meehl (1955), is explained in more detail by Nunnally and Bernstein (1994).

Introducing the topic, Nunnally and Bernstein write: "All sciences, including psychology, are concerned with establishing functional relations between important variables" (Nunnally and Bernstein 1994, 84). They continue:

> To the extent that a variable is abstract and latent rather than concrete and observable ... it is called a "construct." ... A construct reflects a hypothesis (often incompletely formed) that a variety of behaviors will correlate with one another in studies of individual differences and/or will be similarly affected by experimental manipulations. Nearly all theories concern statements about constructs (Nunnally and Bernstein 1994, 85).

Furthermore, they claim: "In general, science's two major concerns are (1) developing measures of individual constructs and (2) finding functional relations between measures of different constructs" (Nunnally and Bernstein 1994, 85). The concept of construct validation, which comes into play when "measuring psychological attributes" (Nunnally and Bernstein 1994, 83), is intended to help scientists reach these goals. Construct validation, according to Nunnally and Bernstein (1994), has three steps. As the two authors put it:

> There are three major aspects of construct validation: (1) specifying the domain of observables related to the construct; (2) determining the extent to which observables tend to measure the same thing, several different things, or many different things from empirical research and statistical analyses; and (3) performing subsequent individual differences studies and/or experiments to determine the extent to which supposed measures of the construct are consistent with "best guesses" about the construct (Nunnally and Bernstein 1994, 86-87).

In what follows, I will discuss the three steps in order. There is no assumption, however, that the three steps take place in this order in the context of actual research.

First, scientists need to identify a class of observable variables that are related to the construct. Nunnally and Bernstein argue that there is no precise method that one can follow in this step. As a result, the scientists must therefore to a great extent rely on intuition and preconceived ideas about how the construct would vary across conditions. Having identified such a class is, nevertheless, a necessary condition for the second step, which is to explore whether the observable variables can be described as measuring the same thing or not. This step is performed by "determining how well the measures of observables 'go together' (intercorrelate) empirically" (Nunnally and Bernstein 1994, 88). This means that the scientist needs to collect data about how the observables vary across conditions, and compute the relevant correlation coefficients. The authors add:

> The results of investigations like those described above lead to one of three conclusions. If all the proposed measures correlate highly with one another, it can be concluded that they all measure much the same thing. If the measures tend to split up into clusters such that the members of a cluster correlate highly with one another and correlate much less with the members of other clusters, they measure a number of *different* things.... A third possibility is that the correlations among the measures all are near zero, so that they measure different things and there is no meaningful construct (Nunnally and Bernstein 1994, 90).

The third step is to show that a set of highly correlated observables in a domain can legitimately be taken to be measures of the construct in which the scientist is interested. In the words of Nunnally and Bernstein: "To the extent that the elements of such a domain [are intercorrelated], *some* construct may be employed to account for the interrelationships, but it is by no means certain that the construct name which motivated the research is appropriate" (Nunnally and Bernstein 1994, 90). To see whether a set of intercorrelated variables can be assumed to be a measure of a given construct (like anxiety, stress or happiness) the scientist needs to explore whether the variables vary across conditions approximate like we would expect degrees of anxiety, stress or happiness to do. Nunnally and Bernstein also propose that construct validity obtains if "the supposed measure(s) of the construct *behave as expected* (Nunnally and Bernstein 1994, 90, italics in original). Presumably, this means that the measure varies across conditions more or less as expected.

The reasoning exhibited by Diener et al. (1999) fits the schematic picture painted by Nunnally and Bernstein (1994) very well. First, Diener et al. have identified a set of "observable" variables – in this case, among other things, spouses' reports and the absence of suicides – that they take to reflect the same construct as self-reported happiness. Second, Diener et al. have explored the degree to which these variables intercorrelate (positively and negatively) with the self-reports. Third, because all these variables (supposedly) correlate positively, Diener et al. conclude that they all are measures of the same construct. Moreover, Diener et al. do check whether the construct "behave as expected." When Diener et al. (1999) argue that self-reports exhibit "appropriate sensitivity to

changing life circumstances" what they mean is that the measure varies across conditions more or less in the expected manner.

Thus, the schematic picture outlined in Kline (1991) and Nunnally and Bernstein (1994) fits well with the argument presented by Diener et al. (1999) in favor of the adequacy of subjective measures of well-being. This strongly supports the contention that psychologists defending subjective measures operate within the psychometric approach to measurement. There are other reasons to believe that psychologists adopt the psychometric approach to measurement too. One is that this approach emerged in the is an outgrowth of personality psychology, which, I have argued, also inspired the development of what we now call subjective measures.

4.4.2 Implications

The contention that psychologists operate within the psychometric approach to measurement sheds some light on the procedure used by psychologists who study well-being, and on various arguments offered for and against the measures. First of all, the contention explains why it comes so natural for psychologists to answer questions about well-being using questionnaires. Questionnaires have been commonly accepted as a tool for mental testing since the very beginning of personality psychology.[69] Moreover, the historical background explains why psychologists think that it is justified to address issues of well-being by using self-reports. It is not that they unquestioningly believe that people are capable of accurately revealing the degree of happiness they enjoy. The use of self-reports is supported by a great deal of previous research, which is interpreted as showing that self-report questionnaires are as valid as any other tool.

[69] This is not to say that there are no differences between using questionnaires to study intelligence and to study happiness, of course.

This history also serves to explain why the economists' approach to welfare measurement tends to strike psychologists as inadequate. For somebody trained in the personality psychology tradition, using people's income as a proxy for their welfare is likely to appear as a hopelessly indirect measure, whether welfare means degree of desire-fulfillment, happiness or satisfaction. This is particularly true since economists make little effort to develop arguments along the lines discussed in this chapter to show their measures of welfare – such as Equivalent Variation – valid as measures of welfare.

Furthermore, the contention that psychologists operate within the psychometric approach helps us identify some assumptions that psychologists do not make. For example, proponents of subjective measures need not assume that individuals in general behave so as to maximize their happiness. Thus, an argument against the claim that people in general maximize happiness is not by itself an argument against subjective measures of well-being. Moreover, the psychologists do not assume that individuals in general are capable of accurately revealing their degree of happiness, only that self-reports are sufficiently correlated with other measures of the construct. While the economists' approach to welfare measurement (the behavioral-representational approach) requires that individuals' choices in fact satisfy the axioms of rational decision, i.e. that these axioms be *true* as a matter of empirical fact, psychologists do not need to assume that individual's revealed happiness always correspond to their actual happiness. Hence, an argument against the claim that people always know how happy or satisfied they really are is not, in itself, an argument against subjective measures of well-being.

Finally, the reference to the psychometric approach tells us something about what a good counterargument would look like. Two possibilities stand out as particularly interesting. Either, the entire approach is misguided. If so, it should be possible to develop a general argument against the psychometric approach across the board, whether for the purposes of measuring intelligence,

optimism/pessimism, or happiness. Or, the psychometric approach is acceptable in principle, but the happiness and satisfaction constructs, and their measures, have not been adequately validated. If so, it should be possible to articulate an argument to the effect that the correlations identified by the psychologists are not strong enough to warrant the their inferences (although other constructs like intelligence may in fact have been adequately validated). To date, the economists criticizing subjective measures have not adopted either one of these strategies.

Of course, the psychologists are aware of problems of the general kind discussed in the previous paragraph. As Diener et al. (1999) 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, 278).[70] In the light of phenomena like these, Norbert Schwarz and Fritz Strack (1999) 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, 61). The exact implications of these findings, and other like them, remain to a great extent undecided. Whether or not the actual correlations identified by the psychological research are strong enough to justify their inferences is worth discussing. Yet, as I hope to have shown, the measurement theoretic argument has no bearing on the issue.

[70] References have been omitted.

4.5 CONCLUSION

In this chapter, I have argued that the economists' measurement theoretic argument is best understood in terms of the Robbins/Hutchison tradition and the behavioral-representational approach to measurement. Furthermore, I have claimed, the argument relies essentially on an empirical premiss, viz. that the choices of real economic agents satisfy the axioms of rational choice theory. Finally, I have suggested, this premiss appears to be, as a matter of fact, false.

The discussion appears to paint a bleak picture for the future of economic measures. However, it seems that Krantz's conclusion is a little too strong. It is possible to maintain a belief in the existence of preferences even in the face of the empirical evidence discussed above. Consistency of choice is not a necessary condition for having preferences. It is quite possible to maintain that people do have preferences, but that preference is related to choice in some less reliable way. This view would have a number of important implications. First, the exact relationship between choices and preferences would be an empirical question, which would need to be settled by scientific research. Second, although choices would still serve as evidence of preference, it would be illegitimate to infer statements about preference s directly from statements about choices.

For the defenders of economic measures of well-being, this strikes me as a reasonable attitude in the face of empirical results like those discussed above. It can still be claimed that choices can be used to develop a rough idea of the level of preference satisfaction of a given individual, and that the amount of resources available to an agent can serve as an indicator of her well-being. Surely, if preferences exist at all, choices can be used as *some* evidence of preferences. The use of choice data as evidence for claims about preference does not, of course, preclude other kinds of evidence. For a long time, Amartya Sen has urged economists to allow e.g. verbal reports (cf. Sen 1982, 71-72).

This move would, however, have some interesting implications for the measurement of welfare. If economists reject the notion that measures of welfare can be based on an observable,

behavioral ordering, the standard arguments for their validity have to be replaced by other arguments. These arguments cannot, obviously, be based on the behavioral-representational approach to measurement. It is difficult to guess what such arguments would look like, or how successful they would be. However, a fair guess is that they would follow the tracks laid down by the psychometric approach. After all, in its terms preferences would be a latent variable, or a "construct," and may be measured according to the methods psychologists have developed for this task.

Either way, I hope to have shown that the psychological measures cannot be dismissed quite as easily as some critics have envisioned. At least, I hope to have demonstrated that the measurement theoretic objection (as I have interpreted it) is not convincing. This does not imply, however, that subjective measures are better than economic ones; as we saw in section 4.4.2, subjective measures are associated with some serious problems too. It does not even imply that degrees of happiness are easier to measure than degrees of preference satisfaction. It does imply, I think, that the choice of measures of well-being is a difficult one, and one which is worth our serious attention. Finally, it suggests that it will not be possible to identify the superior measure of well-being except by facing some purely philosophical questions – such as about the nature of well-being – head on. The conclusion that subjective measures cannot be dismissed on measurement theoretic grounds alone, reinforces the conclusion from the previous chapter, viz. that questions relating to the proper measure of well-being to be used for purposes of public policy will depend on some fundamentally philosophical questions, most prominently about the nature of well-being.

5.0 GENERAL DISCUSSION

Abstract. In this concluding chapter I recapitulate the aims of this dissertation, and examine how the discussion in earlier chapters substantiates the claim made in the introduction. Thus, I demonstrate how the discussion supports the claim that the philosophical foundations of subjective measures are unclear. I also review the evidence in favor of my main thesis, viz. the claim that subjective measures of well-being cannot be shown to be inferior to economic measures quite so easily as some have suggested, but that they nevertheless are associated with serious problems, and that questions about the relative advantage of subjective and economic measures for purposes of public policy will depend on some fundamentally philosophical judgments.

As noted in chapter 1.0, the aim of this dissertation is to explore issues in the philosophical foundations of subjective measures of well-being, and, in particular, how the philosophical foundations of subjective measures differ from those of traditional economic measures. Moreover, the goal is to examine some arguments for and against these measures, and, in particular, arguments that purport to demonstrate the superiority of economic measures (e.g. for purposes of public policy).

As a preliminary, I hope to have substantiated the claim that the philosophical foundations of subjective measures – especially as compared to those of economic measures – are unclear. First, we have seen that it is often unclear what exact account of well-being the psychologists have in mind. In some cases, it appears that authors disagree with themselves about the nature of well-being. Some authors state explicitly that they will not offer a definition. The same is true, to some extent, of economists; it is in fact often unclear what preferences economists think matter to our welfare. Second, it is unclear what approach to measurement is used by economists and psychologists.

Answers to questions like these, I have suggested, helps shed light on the different attempts to measure well-being, and on the arguments offered for and against them.

The discussion, I believe, has substantiated my main thesis, i.e., the claim that subjective measures of well-being cannot be shown to be inferior to economic measures quite so easily as some have suggested, but that they nevertheless are associated with serious problems, and that questions about the relative advantage of subjective and economic measures for purposes of public policy will depend on some fundamentally philosophical judgments, e.g. about the nature of well-being and the legitimate goals for public policy. In what follows I will recapitulate the evidence brought to bear on each of the different components of this thesis.

As we have seen, a number of critics – often economists who are comfortable using traditional economic measures – have suggested that subjective measures of well-being can be dismissed relatively easily. These critics include Anders Åslund, whose argument appears to be based on the contention that hedonism is an inadequate account of well-being (see section 3.6). They also include Wilfred Beckerman, who suggested that mental state accounts like those used by psychologists do not permit the development of adequate measures, unlike preference satisfaction accounts of well-being (see section 4.1).

Yet, as we have also seen, these arguments do not in fact show that subjective measures are inferior to traditional economic measures for purposes of public policy. In brief, actual preference satisfaction accounts are just as implausible as narrow hedonism and related accounts. While it is possible to defend economic measures as indirect indicators of more sophisticated accounts of well-being, it is not clear that they are superior as indirect measures either. Similarly, the argument that preference satisfaction but not happiness can be measures fails because it is based on a false empirical premiss. The critics of subjective measures have not shown that there is a difference between economic and subjective measures in this regard.

While subjective measures cannot be dismissed quite as easily as some have suggested, they remain associate with serious issues. The first relates to the underlying account of well-being. As we have already seen, the account of well-being used by proponents of subjective measures are implausible, and it is unclear whether subjective measures can be argued to represent well-being on some other, more sophisticated account. Moreover, in the context of public policy, it can be argued that it is a mistake to pay too much attention to well-being. It can be argued that governments should be concerned with promoting the opportunities open to people, or the material and other requisites for a life of well-being, or that which *should* make people happy rather than what in fact makes them happy.[71] Finally, there are some serious problems associated with the measurement of mental states like happiness, including the fact that answers to questions about happiness and the like are strongly context- and frame-dependent.

This does not mean that efforts to develop subjective measures are misplaced, or that the whole "science of happiness" is misguided. Subjective measures may be quite interesting as measures of happiness, satisfaction, and the like, rather than measures of well-being, just like measures of GDP can be quite interesting as measures of production rather than welfare. This realization may be implicit in the occasional tendency to refer to measures of happiness and the like as "measures of subjective well-being" or "measures of psychological well-being" rather than as "subjective measures of well-being." Moreover, as indicated in section 3.7, measures of happiness may also serve as indirect measures of well-being (properly understood).

Anyway, it appears that questions related to the relative advantages of traditional economic and subjective measures cannot be answered without addressing some fundamentally philosophical issues. The most prominent issues concern the nature of well-being, that is, whether well-being is a matter of mental states of some kind, preference satisfaction, or something else. Many of the

[71] See Sen (1987) and Rescher (1972) for discussions of these views.

arguments offered both for and against subjective measures can be interpreted as attempts to sidestep such questions Åslund and Beckerman quite clearly wish to dismiss subjective measures of well-being without having to address questions like the nature of well-being. The same thing can be said for psychologists who avoid explicit discussion of the relationship between well-being and happiness, proceeding as if it were uncontroversially true that the two are the same. Yet, at the end of the day, many arguments for and against various measures of well-being will depend critically on considerations of a distinctly philosophical nature. Thus, the question about the appropriate measures of well-being for a given purpose will have to be settled on the basis of a variety of grounds, economic, psychological, and philosophical.

BIBLIOGRAPHY

Adams, David L. (1971) 'Correlates of Satisfaction among the Elderly,' *Gerontologist* 11: 64-68.

Ahlheim, Michael (1998) 'Measures of Economic Welfare,' in Barberà et al. (1998), pp. 483-568.

Allen, Mary J. and Wendy M. Yen (1979) *Introduction to Measurement Theory* (Monterey: Brooks/Cole).

Allport, Gordon W. (1937) *Personality: A psychological interpretation* (New York: H. Holt, Rinehart & Winston).

Andrews, Frank M. (1989) 'The Evolution of a Movement,' *Journal of Public Policy* 9: 401-405.

Andrews, Frank M. and John P. Robinson (1991) 'Measures of Subjective Well-Being,' in John P. Robinson, Phillip R. Shaver, and Lawrence S. Wrightsman (Eds.) *Measures of Personality and Social Psychological Attitudes* (San Diego: Academic Press), pp. 61-114.

Angner, Erik (in progress) 'Economists as Experts: Overconfidence in theory and practice.'

Argyle, Michael (1999) 'Causes and Correlates of Happiness,' in Kahneman et al. (1999), pp. 354-373.

Arrow, Kenneth J., Enrico Colombatto, Mark Perlman, and Christian Schmidt (Eds.) (1996) *The Rational Foundations of Economic Behaviour: Proceedings from the IEA Conference held in Turin, Italy* (New York: St. Martin's).

Arrow, Kenneth and Tibor Scitovsky (Eds.) (1969) *Readings in Welfare Economics* (Homewood: Richard D. Irwin).

Ashraf, Nava, Colin F. Camerer, and George Loewenstein (manuscript) 'Adam Smith, Behavioral Economist,' 20 May 2004.

Åslund, Anders (1994) 'Russia's Success Story,' *Foreign Affairs* 73: 58-71.

Åslund, Anders (2001) 'The Myth of Output Collapse after Communism,' Carnegie Working Paper, Carnegie Endowment for International Peace.

Barberà, Salvador, Peter J. Hammond, and Christian Seidl (1998) *Handbook of Utility Theory*, Vol. I. (Dordrecht: Kluwer).

Bateman, Bradley W. (2001) 'Make a Righteous Number: Social surveys, the men and religion forward movement, and quantification in American economics,' in Klein and Morgan (2001), pp. 57-85.

Beckerman, Wilfred (1987) 'National Income,' in John Eatwell, Murray Milgate, Peter Newman (Eds.) *The Palgrave Dictionary of Economics*, Vol. III (New York: Stockton Press), pp. 590-592.

Beckerman, Wilfred (1974) *Two Cheers for the Affluent Society: A spirited defense of economic growth* (New York: St. Martin's Press).

Bentham, Jeremy (1996 [1789]) *Introduction to the principles of morals and legislation*, (Eds.) J. H. Burns and H. L. A. Hart (Oxford: Clarendon).

Binmore, Ken (1994) *Game Theory and the Social Contract*, Vol. I (Cambridge: MIT Press).

Blackorby, Charles and David Donaldson (1990) 'A Review Article: The case against the use of the sum of compensating variations in cost-benefit analysis,' *Canadian Journal of Economics* 23: 471-494.

Bogen, James and James Woodward (1988) 'Saving the Phenomena,' *The Philosophical Review* 97: 303-352.

Bradburn, Norman M. (1969) *The Structure of Psychological Well-Being* (Chicago: Aldine).

Bradburn, Norman M. and David Caplovitz (1965) *Reports on Happiness: A pilot study of behavior related to mental health* (Chicago: Aldine).

Bridgman, Percy W. (1927) *The Logic of Modern Physics* (New York: Macmillan).

Burgess, Ernest W. and Leonard S. Cottrell (1939) *Predicting Success or Failure in Marriage* (New York: Prentice-Hall).

Campbell, Angus (1976) 'Subjective Measures of Well-Being,' *American Psychologist* 31: 117-124.

Campbell, Angus, Philip E. Converse, and Willard L. Rodgers (1976) *The Quality of American Life: Perceptions, evaluations, and satisfactions* (New York: Russell Sage).

Cantril, Hadley (1965) *The Pattern of Human Concerns* (New Brunswick: Rutgers University Press).

Carley, Michael (1981) *Social Measurement and Social Indicators: Issues of policy and theory* (Boston: George Allen & Unwin).

Cavan, Ruth Shonle, Ernest W. Burgess, Robert J. Havighurst, and Herbert Goldhamer (1949) *Personal Adjustment in Old Age* (Chicago: Science Research Associates).

Chappell, Tim and Roger Crisp (1998) 'Utilitarianism,' in Edward Craig (Ed.) *The Routledge Encyclopedia of Philosophy*, Vol. 9 (London: Routledge), pp. 551-557.

Chipman, John S. and James C. Moore (1976) 'Why an Increase in GNP Need Not Imply an Improvement in Potential Welfare,' *Kyklos* 29: 391-418.

Daneshkhu, Scheherazade (2005) 'The "dismal science" turns its attention to question of happiness,' *Financial Times*, 10 January.

Davis, Katharine Bement (1929) *Factors in the Sex Life of Twenty-Two Hundred Women* (New York: Harper & Brothers).

Dawes, Robyn M. (1988) *Rational Choice in an Uncertain World* (Fort Worth: Harcourt Brace Jovanovich).

Denison, Edward F. (1971) 'Welfare Measurement and the GNP,' *Survey of Current Business* 51: 13-39.

Diener, Ed (1984) 'Subjective Well-Being,' *Psychological Bulletin* 95: 542-575.

Diener, Ed and Robert Biswas-Diener (2002) 'Will Money Increase Subjective Well-Being? A literature review and guide to needed research,' *Social Indicators Research* 57: 119-169.

Diener, Ed, Richard E. Lucas, and Shigehiro Oishi (2002) 'Subjective Well-Being: The science of happiness and life satisfaction,' in Snyder and Shane Lopez (2002), pp. 63-73.

Diener, Ed and Eunkook Suh (1997) 'Measuring Quality of Life: Economic, social, and subjective indicators,' *Social Indicators Research* 40: 189-216.

Diener, Ed, Eunkook M. Suh, Richard E. Lucas, and Heidi L. Smith (1999) 'Subjective Well-Being: Three decades of progress,' *Psychological Bulletin* 125: 276-303.

Duncan, Otis Dudley (1969) *Toward Social Reporting: Next steps* (New York: Russell Sage).

Dupuit, Jules (1969) 'On the Measurement of Public Works,' in Arrow and Scitovsky (1969), pp. 255-283. Originally published as 'De la mesure de l'utilité des travaux publics' in *Annales des Ponts et Chaussées* 8 (1844).

Easterbrook, Gregg (2003) *The Progress Paradox: How life gets better while people feel worse* (New York: Random House).

Easterlin, Richard E. (1974) 'Does Economic Growth Improve the Human Lot? Some empirical evidence,' in Paul A. David and Melvin W. Reder (Eds.) *Nations and Households in Economic Growth: Essays in honor of Moses Abramovitz* (New York: Academic Press), pp. 89-125. Reprinted in Easterlin (2002), pp. 5-41.

Easterlin, Richard E. (2001)'Income and Happiness: Towards a unified theory,' *The Economic Journal* 111: 465-484. Reprinted in Easterlin (2002), pp. 206-225.

Easterlin, Richard E. (Ed.) (2002) *Happiness in Economics* (Cheltenham: Elgar).

Elster, Jon and Aanund Hylland (1986) *Foundations of Social Choice Theory* (Cambridge: Cambridge University Press).

Fehige, Christoph and Ulla Wessels (Eds.) (1998) *Preferences* (Berlin: de Gruyter).

Flügel, J. C. (1925) 'A Quantitative Study of Feeling and Emotion in Everyday Life,' in *The British Journal of Psychology* 15: 318-355.

Frederick, Shane and George Loewenstein (1999) 'Hedonic Adaptation,' in Kahneman et al. (1999), pp. 302-329.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002) 'Time Discounting and Time Preference: A critical review,' *Journal of Economic Literature* 40: 351-401. Reprinted in Loewenstein et al. (2003), pp. 13-86.

Frey, Bruno S. and Alois Stutzer (2000) 'Maximizing Happiness?,' *German Economic Review* 1: 145-167.

Frey, Bruno S. and Alois Stutzer (2002) 'What Can Economists Learn from Happiness Research?,' *Journal of Economic Literature* 40: 402-435.

Ghiselli, Edwin E., John P. Campbell, and Sheldon Zedeck (1981) *Measurement Theory for the Behavioral Sciences* (San Francisco: W. H. Freeman).

Gillham, Jane E. and Martin E. P. Seligman (1999) 'Footsteps on the road to a positive psychology,' *Behaviour and Research Therapy* 37: S163-S173.

Goldings, Herbert Jeremy (1954) 'On the Avowal and Projection of Happiness,' *Journal of Personality* 23: 30-47.

Goodin, Robert E. (1986) 'Laundering Preferences,' in Elster and Hylland (1986), pp. 75-101.

Gordon, H. Scott (1991) *The History and Philosophy of Social Science* (London: Routledge).

Griffin, J. P. (1998) 'Happiness,' in Edward Craig (Ed.) *The Routledge Encyclopedia of Philosophy*, Vol. 4 (London: Routledge), pp. 226-229.

Griffin, James (1982) 'Modern Utilitarianism,' *Revue Internationale de Philosophie* 36: 331-375.

Griffin, James (1986) *Well-Being: Its meaning, measurement and moral importance* (Oxford: Clarendon).

Griffin, James (1996) *Value Judgement: Improving our ethical beliefs* (Oxford: Oxford University Press).

Gurin, Gerald, Joseph Veroff, and Sheila Feld (1960) *Americans View Their Mental Health: A nationwide interview survey* (New York: Basic Books).

Hamilton, Gilbert V. (1929) *A Research in Marriage* (New York: Albert & Charles Boni).

Hare, R. M. (1982) 'Ethical Theory and Utilitarianism,' in Sen and Williams (1982), pp. 23-38.

Harsanyi, John C. (1982) 'Morality and the theory of rational behavior,' in Sen and Williams (1982), pp. 39-62. Originally published in *Social Research* 44 (1977): 623-656.

Hart, Hornell (1940) *Chart for Happiness* (New York: Macmillan).

Hartmann, George W. (1934) 'Personality Traits Associated with Variations in Happiness,' *The Journal of Abnormal and Social Psychology* 29: 202-212.

Hausman, Daniel M. (1993) 'Liberalism, Welfare Economics, and Freedom,' *Social Philosophy and Policy* 10: 172-197.

Hausman, Daniel M (2000) 'Revealed Preference, Belief, and Game Theory,' *Economics and Philosophy* 16: 99-115.

Hausman, Daniel M. and Michael S. McPherson (1996) *Economic Analysis and Moral Philosophy* (Cambridge: Cambridge University Press).

Hausman, Daniel M. and Michael S. McPherson (1997) 'Beware of Economists Bearing Advice,' *Policy Options* 18: 16-19.

Havighurst, Robert J. (1957) 'The Social Competence of Middle-Aged People,' *Genetic Psychology Monographs* 56: 297-375.

Haybron, Daniel M. (2000) 'Two Philosophical Problems in the Study of Happiness,' *Journal of Happiness Studies* 1: 207-225.

Haybron, Daniel M. (2003) 'What Do We Want from a Theory of Happiness?,' *Metaphilosophy* 34: 305-329.

Hempel, Carl G. (1965) 'A Logical Appraisal of Operationism,' in *Aspects of Scientific Explanation* (New York: Free Press), pp. 123-133. Originally published in *Scientific Monthly* 79 (1954): 215-220.

Hicks, John R. (1939) 'The Foundations of Welfare Economics,' *Economic Journal* 49: 696-712.

Hicks, John R. (1943) 'The Four Consumer's Surpluses,' *Review of Economic Studies* 11: 31-41.

Hicks, John R. (1956) *A Revision of Demand Theory* (Oxford: Clarendon).

Hicks, John R. (1975) 'The Scope and Status of Welfare Economics,' *Oxford Economic Papers* 27: 307-326.

Hutchison, T. W. (1938) *The Significance and Basic Postulates of Economic Theory* (London: Macmillan).

Johansson, Per-Olov (1991) *An Introduction to Modern Welfare Economics* (Cambridge: Cambridge University Press).

Johnson, J. A. (2001) 'Personality Psychology: Methods,' in Smelser and Baltes (2001), Vol. 16, pp. 11313-11317.

Just, Richard E., Darrell L. Hueth, and Andrew Schmitz (1982) *Applied Welfare Economics and Public Policy* (Englewood Cliffs: Prentice-Hall).

Kahneman, Daniel (1999) 'Objective Happiness,' in Kahneman et al. (1999), pp. 3-25.

Kahneman, Daniel (2000a) 'Experienced Utility and Objective Happiness,' in Kahneman and Tversky (2000), pp. 673-692.

Kahneman, Daniel (2000b) 'Evaluation by Moments,' in Kahneman and Tversky (2000), pp. 693-708.

Kahneman, Daniel (2000c) 'New Challenges to the Rationality Assumption,' in Kahneman and Tversky (2000), pp. 758-774.

Kahneman, Daniel, Ed Diener, and Norbert Schwarz (Eds.) (1999) *Well-Being: The foundations of hedonic psychology* (New York: Russell Sage).

Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur Stone (2004) 'Toward National Well-Being Accounts,' *American Economic Review* 94: 429-434.

Kahneman, Daniel, Ilana Ritov, and David Schkade (1999) 'Economic Preferences or Attitude Expressions? An analysis of dollar responses to public issues,' *Journal of Risk and Uncertainty* 19: 203-235.

Kahneman, Daniel, Paul Slovic, and Amos Tversky (1982) *Judgment under Uncertainty: Heuristics and biases* (Cambridge: Cambridge University Press).

Kahneman, Daniel and Amos Tversky (Eds.) (2000) *Choices, Values and Frames* (Cambridge: Cambridge University Press).

Kasser, Tim (2002) *The High Price of Materialism* (Cambridge: MIT Press).

Keller, Simon (2004) 'Welfare and the Achievement of Goals,' *Philosophical Studies* 121: 27-41.

Klein, Judy L. and Mary S. Morgan (2001) *The Age of Economic Measurement* (Durham: Duke University Press).

Kline, Paul (1998) *The New Psychometrics: Science, psychology and measurement* (London: Routledge).

Korsgaard, Christine M. (1993) 'G. A. Cohen: Equality of What? On Welfare, Goods and Capabilities. Amartya Sen: Capability and Well-Being,' in Nussbaum and Sen (1993), pp. 54-61.

Krantz, David H. (1972) 'Measurement Structures and Psychological Laws,' *Science* 175: 1427-1435.

Krantz, David H. (1991) 'From Indices to Mappings: The representational approach to measurement,' in Donald R. Brown and J. E. Keith Smith (Eds.) *Frontiers of Mathematical Psychology: Essays in honor of Clyde Coombs* (New York: Springer), pp. 1-52.

Krantz, David H., R. Duncan Luce, Patrick Suppes, and Amos Tversky (1971) *Foundations of Measurement Vol. I: Additive and polynomial representations* (New York: Academic Press).

Kuhlen, R. G. (1948) 'Age Trends in Adjustment during the Adult Years as Reflected in Happiness Ratings [Abstract],' *American Psychologist* 3: 307.

Kutner, Bernard, David Fanshel, Alice M. Togo, and Thomas S. Langner (1956) *Five hundred over sixty: A community survey on aging* (New York: Russell Sage).

Kyburg, Henry E. (1984) *Theory and Measurement* (Cambridge: Cambridge University Press).

Lane, Robert E. (2000) *The Loss of Happiness in Market Democracies* (New Haven: Yale University Press).

Larsen, Randy, Ed Diener, and Robert Emmons (1985) 'An Evaluation of Subjective Well-Being Measures,' *Social Indicators Research* 17: 1-17.

Larson, Reed (1978) 'Thirty Years of Research on the Subjective Well-Being of Older Americans,' *Journal of Gerontology* 33: 109-125.

Lawton, M. Powell (1975) 'The Philadelphia Geriatric Center Morale Scale: A revision,' *Journal of Gerontology* 30: 85-89.

Layard, Richard (2005) *Happiness: Lessons from a new science* (New York: Penguin Press).

Lebo, Dell (1953) 'Some Factors Said to Make for Happiness in Old Age,' *Journal of Clinical Psychology* 9: 385-387.

LeDoux, Joseph (1996) *The Emotional Brain: The mysterious underpinnings of emotional life* (New York: Touchstone).

LeDoux, Joseph and Jorge Armony (1999) 'Can Neurobiology Tell Us Anything About Human Feeling?,' in Kahneman et al. (1999), pp. 489-499.

Levi, Isaac (1997) *The Covenant of Reason: Rationality and the commitments of thought* (Cambridge: Cambridge University Press).

Levy, Ed (2001) 'Quantification, Mandated Science and Judgment,' *Studies in History and Philosophy of Science* 32: 723-737.

Lichtenstein, Sarah and Paul Slovic (1971) 'Reversals of Preference Between Bids and Choices in Gambling Decisions,' *Journal of Experimental Psychology* 89: 46-55.

Little, Ian M. D. (1949) 'A Reformulation of the Theory of Consumer's Behaviour,' *Oxford Economic Papers*, 1: 90-99.

Loewenstein, George (1999) 'Because It Is There: The challenge of mountaineering… for utility theory,' *Kyklos* 52: 315-344.

Loewenstein, George and Erik Angner (2003) 'Predicting and Indulging Changing Preferences,' in Loewenstein et al. (2003), pp. 351-391.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin (working paper) 'Projection bias in predicting future utility'.

Loewenstein, George, Daniel Read, and Roy Baumeister (Eds.) (2003) *Time and Decision: Economic and psychological perspectives on intertemporal choice* (New York: Russell Sage Foundation Press).

Loewenstein, George and David Schkade (1999) 'Wouldn't It Be Nice? Predicting future feelings,' in Kahneman et al. (1999), pp. 85-105.

Loewenstein, George and Richard H. Thaler (1989) 'Anomalies: Intertemporal Choice,' *Journal of Economic Perspectives* 3: 181-193.

Lucas, Richard E., Ed Diener, and Eunkook Suh (1996) 'Discriminant Validity of Well-Being Measures,' *Journal of Personality and Social Psychology* 71: 616-628.

Luce, R. Duncan, David H. Krantz, Patrick Suppes, and Amos Tversky (1990) *Foundations of Measurement Vol. III: Representation, axiomatization, and invariance* (New York: Academic Press).

Machamer, Peter (1998) 'Individual and Other Person Morality: A plea for an emotional response to ethical problems,' in Axel Wüstehube and Michael Quante (Eds.) *Pragmatic Idealism: Critical essays on Nicholas Rescher's system of pragmatic idealism,* in *Poznan Studies in the Philosophy of the Sciences and the Humanities* 64: 73-84.

Mankiw, N. Gregory (2001) *Principles of Microeconomics*, Second Ed. (Fort Worth: Harcourt College Publishers).

Marshall, Alfred (1948) *Principles of Economics: An introductory volume*, Eight Ed. (New York: Macmillan). First Ed. (London: Macmillan, 1890).

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995) *Microeconomic Theory* (Oxford: Oxford University Press).

McAdams, O. P. (2001) 'Personality Psychology,' in Smelser and Baltes (2001), Vol. 16, pp. 11308-11313.

Mill, John Stuart (1998 [1861]) *Utilitarianism*, (Ed.) R. Crisp (Oxford: Clarendon).

Mongin, Philippe and Claude d'Aspremont (1998) 'Utility Theory and Ethics,' in Barberà et al. (1998), pp. 371-481.

Moore, Andrew and Roger Crisp (1996) 'Welfarism in Moral Theory,' *Australasian Journal of Philosophy* 74: 598-613.

Morgan, Mary S. (2001) 'Making Measuring Instruments,' in Klein and Morgan (2001), pp. 235-251.

Myers, David G. (1993) *The Pursuit of Happiness: Who is happy and why?* (New York: William Morrow).

Myerson, Abraham (1918) 'Eupathics—A Program for Mental Hygiene,' *Journal of Abnormal and Social Psychology* 12: 343-347.

Neugarten, Bernice L., Robert J. Havighurst, and Sheldon S. Tobin (1961) 'The Measurement of Life Satisfaction,' *Journal of Gerontology* 16:134-143.

von Neumann, John and Oskar Morgenstern (1944) *Theory of Games and Economic Behavior* (Princeton: Princeton University Press).

Nozick, Robert (1974) *Anarchy, State and Utopia* (Oxford: Blackwell).

Nunnally, Jum C. (1978) *Psychometric Theory*, Second Ed. (New York: McGraw-Hill).

Nussbaum, Martha and Amartya Sen (Eds.) (1993) *The Quality of Life* (Oxford: Oxford University Press).

Oswald, Andrew J. (1997) 'Happiness and Economic Performance,' *The Economic Journal* 107: 1815-1831. Reprinted in Easterlin (2002), pp. 149- 165.

Parfit, Derek (1984) *Reasons and Persons* (Oxford: Clarendon).

Pigou, A. C. (1960) *The Economics of Welfare*, Fourth Ed. (London: Macmillan). First Ed. (London: Macmillan, 1920).

Plant, Raymond (1991) *Modern Political Thought* (Oxford: Blackwell).

Plott, Charles (1996) 'Rational Individual Behavior in Markets and Social Choice Processes: The discovered preference hypothesis,' in Arrow et al. (1996), pp. 225-250.

Pollak, Otto (1948) *Social Adjustment in Old Age: A research planning report* (New York: Social Science Research Council).

Porter, Theodore M. (1994) 'Making Things Quantitative,' *Science in Context* 7:389-407.

Porter, Theodore M. (1995) *Trust in Numbers: The pursuit of objectivity in science and public life* (Princeton: Princeton University Press).

Porter, Theodore M. (2001) 'On the Virtues and Disadvantages of Quantification for Democratic Life,' *Studies in History and Philosophy of Science* 32: 739-747.

Qizilbash, Mozaffar (1998) 'The Concept of Well-Being,' *Economics and Philosophy* 14: 51-73.

Ramsey, Frank P. (1931) 'Truth and Probability,' in R. B. Braithwaite (Ed.) *The Foundations of Mathematics* (London: Routledge and Kegan Paul), pp. 156-198. Reprinted in Peter Gärdenfors and Nils-Eric Sahlin (Eds.) *Decision, Probability, and Utility: Selected readings* (Cambridge: Cambridge University Press, 1988), pp. 19-47.

Rawls, John (1999) *A Theory of Justice*, Revised Ed. (Cambridge: Belknap).

Redelmeier Donald A. and Daniel Kahneman (1996) 'Patients' Memories of Painful Medical Treatments: Real-time and retrospective evaluations of two minimally invasive procedures,' *Pain* 66: 3-8.

Rescher, Nicholas (1972) *Welfare: The social issues in philosophical perspective* (Pittsburgh: University of Pittsburgh Press).

Rescher, Nicholas (1980) *Unpopular Essays on Technological Progress* (Pittsburgh: University of Pittsburgh Press).

Robbins, Lionel C. (1984) *An Essay on the Nature and Significance of Economic Science*, Third Ed. (New York: New York University Press). First Ed. (London: Macmillan, 1932).

Roberts, Fred S. (1979) *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences* (Reading: Addison-Wesley).

Rose, Arnold M. (1955) 'Factors Associated with the Life Satisfaction of Middle-Class, Middle-Aged Persons,' *Marriage and Family Living* 17: 15-19.

Rosenberg, Alexander (1981) 'A Skeptical History of Microeconomic Theory,' in Joseph C. Pitt (Ed.) *Philosophy in Economics* (Dordrecht: D. Reidel), pp. 47-61.

Rosenberg, Alexander (1992) *Economics – Mathematical Politics or Science of Diminishing Returns?* (Chicago: University of Chicago Press).

Rossi, Robert J. and Kevin J. Gilmartin (1980) *The Handbook of Social Indicators: Sources, characteristics, and analysis* (New York: Garland).

Rutherford, Donald (2002) *Routledge Dictionary of Economics*, Second Ed. (London: Routledge).

Sailer, Randolph C. (1931) 'Happiness Self-Estimates of Young Men,' *Teachers College Columbia University Contributions to Education*, No. 467.

Samuelson, Paul A. (1938a) 'A Note on the Pure Theory of Consumer's Behaviour.' *Economica,* 5:61-71.

Samuelson, Paul A. (1938b) 'A Note on the Pure Theory of Consumer's Behaviour: An Addendum.' *Economica,* 5:353-354.

Samuelson, Paul A. (1947) *Foundations of Economic Analysis* (Cambridge: Harvard University Press).

Samuelson, Paul A. (1948) 'Consumption Theory in Terms of Revealed Preference,' *Economica,* 15: 243-253.

Samuelson, Paul A. (1952) 'Probability, Utility, and the Independence Axiom,' *Econometrica* 20: 670-678.

Savage, Leonard J. (1972) *Foundations of Statistics*, Second Revised Ed. (New York: Dover). First Ed. (New York: John Wiley and Sons, 1954).

Scanlon, Thomas M. (1975) 'Preference and Urgency,' *The Journal of Philosophy* 72: 655-669.

Scanlon, Thomas M. (1993) 'Value, Desire, and Quality of Life,' in Nussbaum and Sen (1993), pp. 185-200.

Scanlon, Thomas M. (1998) 'The Status of Well-Being,' in Grethe B. Peterson (Ed.) *The Tanner Lectures on Human Values*, Vol. 19 (Salt Lake City: The University of Utah Press), pp. 91-143.

Schwarz, Norbert and Fritz Strack (1999) 'Reports of Subjective Well-Being: Judgmental processes and their methodological implications,' in Kahneman et al. (1999), pp. 61-84.

Scitovsky, Tibor (1951) 'The State of Welfare Economics,' *American Economic Review* 41: 303-315.

Scott, Dana and Patrick Suppes (1958) 'Foundational Aspects of Theories of Measurement,' *The Journal of Symbolic Logic* 23: 113-128.

Seligman, Martin E. P. (2002) *Authentic Happiness: Using the new positive psychology to realize your potential for lasting fulfillment* (New York: Free Press).

Seligman, Martin E. P. and Mihaly Csikszentmihalyi (2000) 'Positive Psychology,' *American Psychologist* 55: 5-14.

Sen, Amartya (1979a) 'Utilitarianism and Welfarism,' *The Journal of Philosophy* 76: 463-489.

Sen, Amartya (1979b) 'The Welfare Basis of Real Income Comparisons: A survey,' *Journal of Economic Literature* 17: 1-45.

Sen, Amartya (1982) *Choice, Welfare and Measurement* (Cambridge: Harvard University Press).

Sen, Amartya (1985) 'Well-Being, Agency and Freedom: The Dewey Lectures 1984,' *The Journal of Philosophy* 82: 169-221.

Sen, Amartya (1987) *Commodities and Capabilities* (New Delhi: Oxford University Press).

Sen, Amartya (1988) 'The Concept of Development,' in H. Chenery and T. N. Srinivasan (Eds.) *Handbook of Development Economics,* Vol. I (New York: Elsevier), pp. 9-26.

Sen, Amartya (1993) 'Capabilities and Well-Being,' in Nussbaum and Sen (1993), pp. 30-53.

Sen, Amartya (1999) *Development as Freedom* (New York: Anchor Books).

Sen, Amartya (2002) *Rationality and Freedom* (Cambridge: Belknap).

Sen, Amartya and Bernard Williams (Eds.) (1982) *Utilitarianism and Beyond* (Cambridge: Cambridge University Press).

Slesnick, Daniel T. (1998) 'Empirical Approaches to the Measurement of Welfare,' *Journal of Economic Literature* 36: 2108-2165.

Slovic, Paul (1995) 'The Construction of Preferences,' *American Psychologist* 50: 364-371.

Smart, J. J. C. (1973) 'An Outline of a System of Utilitarian Ethics,' in J. J. C. Smart and Bernard Williams (Eds.) *Utilitarianism: For and against* (Cambridge: Cambridge University Press).

Smelser, Neil J. and Paul B. Baltes (Eds.) (2001) *International Encyclopedia of the Social and Behavioral Sciences* (Amsterdam: Elsevier).

Snyder, C. R. and Shane J. Lopez (2002) *Handbook of Positive Psychology* (Oxford: Oxford University Press).

Sobel, David (1994) 'Full Information Accounts of Well-Being,' *Ethics* 104: 784-810.

Sobel, David (1998) 'Well-Being as the Object of Moral Consideration,' *Economics and Philosophy* 14: 249-281.

Steiner, Jacob E., Dieter Glaser, Maria E. Hawilo, and Kent C. Berridge (2001) 'Comparative Expression of Hedonic Impact: Affective reactions to taste by human infants and other primates,' *Neuroscience and Biobehavioral Reviews* 25: 53-74.

Strack, Fritz, Michael Argyle, and Norbert Schwarz (1991) *Subjective Well-Being: An interdisciplinary perspective* (Oxford: Pergamon).

Sugden, Robert (1992) 'Individual Choice: Consumer Theory,' in Shaun Hargreaves Heap, Martin Hollis, Bruce Lyons, Robert Sugden, Albert Weale (Eds.) *The Theory of Choice: A critical guide* (Oxford: Blackwell), pp. 26-35.

Sumner, L. W. (1996) *Welfare, Happiness, and Ethics* (Oxford: Clarendon).

Suppes, Patrick, David H. Krantz, R. Duncan Luce, and Amos Tversky (1989) *Foundations of Measurement Vol. II: Geometrical, threshold, and probabilistic representations* (New York: Academic Press).

Suppes, Patrick and Joseph L. Zinnes (1963) 'Basic Measurement Theory,' in R. Duncan Luce, Robert R. Bush, and Eugene Galanter (Eds.) *Handbook of Mathematical Psychology*, Vol. I (New York: Wiley), pp. 1-76.

Symonds, Percival M. (1937) 'Happiness as Related to Problems and Interests,' *The Journal of Educational Psychology* 28: 290-294.

Terman, Lewis M. (1938) *Psychological Factors in Marital Happiness* (New York: McGraw-Hill).

Tinbergen, Jan (1991) 'On the Measurement of Welfare,' *Journal of Econometrics* 50: 7-13.

Tversky, Amos (1996) 'Rational Theory and Constructive Choice,' in Arrow et al. (1996), pp. 185-197.

Tversky, Amos and Daniel Kahneman (1982) 'Judgment Under Uncertainty: Heuristics and biases,' in Kahneman et al. (1982), pp. 3-20. Originally published in *Science* 185 (1974): 1124-1131.

United Nations Development Programme (2004) *Human Development Report 2004: Cultural liberty in today's diverse world* (New York: United Nations Development Programme).

Veenhoven, Ruut (1984) *Conditions of Happiness* (Dordrecht: Reidel).

Velleman, J. David (1991) 'Well-Being and Time,' *Pacific Philosophical Quarterly* 72: 48-77.

Waldfogel, Joel (1993) 'The Deadweight Loss of Christmas,' *American Economic Review* 83: 1328-1336.

Watson, Goodwin (1930) 'Happiness Among Adult Students of Education,' *The Journal of Educational Psychology* 21: 79-109.

Watson, John B. (1924) *Behaviorism* (New York: W. W. Norton).

Weale, Albert (1998) 'Welfare,' in Edward Craig (Ed.) *The Routledge Encyclopedia of Philosophy*, Vol. 9 (London: Routledge), pp. 702-706.

Wessman, Alden E. and David F. Ricks (1966) *Mood and Personality* (New York: Holt, Rinehart and Winston).

Willig, Robert D. (1976) 'Consumer's Surplus Without Apology,' *American Economic Review* 66: 589-597.

Wilson, Timothy D. and Daniel T. Gilbert (working paper) 'Affective Forecasting.'

Wilson, Warner (1967) 'Correlates of Avowed Happiness,' *Psychological Bulletin* 67: 294-306.

Winn, Philip (Ed.) (2001) *Dictionary of Biological Psychology* (London: Routledge).

Wohlgemuth, A. (1919) Pleasure–Unpleasure: An experimental investigation on the feeling-elements, in The British Journal of Psychology Monograph Supplements, Vol. VI (Cambridge: Cambridge University Press).

World Bank (2004) *World Development Indicators Online* (Washington D.C.).