Health Care Anal (2008) 16:97–113
DOI 10.1007/s10728-007-0066-4

ORIGINAL ARTICLE

Well-Being and Health

Greg Bognar

Published online: 15 November 2007

Abstract One way of evaluating health is in terms of its impact on well-being. It has been shown, however, that evaluating health this way runs into difficulties, since health and other aspects of well-being are not separable. At the same time, the practical implications of the inseparability problem remain unclear. This paper assesses these implications by considering the relations between theories, components, and indicators of well-being.

Keywords Well-being · Health · Indicators · Inseparability · Preferences

The Inseparability of Health and Well-Being

In order to make medical resource allocation and prioritization decisions, we need to be able to evaluate health. By "evaluating health," I mean measuring the value of health. On the view I shall be concerned with, this is understood as measuring how good (or bad) a person's health state is for that person—that is, how a person's health contributes to that person's well-being.

Measuring the *value* of health is different from measuring the *amount* of health. The objective of the latter is to assess health without its impact on functioning, ability, or quality of life, taken in a broad sense. It is concerned with disease and impairment. While it seems uncontroversial to say that a person is healthier without a disease than with it, such comparisons cannot be taken far. It seems arbitrary to say, for example, that a person is less healthy (has less health) if she has poor eyesight than if she is hard of hearing, unless we mean that having good eyesight is more *valuable* than having good hearing. In order to compare health states, we inevitably have to appeal to their value.

G. Bognar (✉)
Program in Ethics & Health, Harvard University, 641 Huntington Ave, Boston, MA 02115, USA
e-mail: greg_bognar@hms.harvard.edu

Measuring the amount of health and evaluating health can thus come apart: a person may be in poor health at some time, but her condition may be asymptomatic at that time. Since it does not have any impact, her present well-being is unaffected. More often, a person's health has an impact on her well-being, but her well-being is affected by other factors as well. For instance, people with chronic illnesses may adapt to their changed circumstances by adjusting their goals and activities. At least in some of these cases—especially if the new goals and activities are worthwhile and also experienced as fulfilling—the impact of ill-health on the person's well-being is mitigated by successful adaptation.

There are other ways of measuring the value of health. For instance, health may be evaluated in terms of its consequences for productivity—a way of measuring its contribution to people's "social value." Another alternative is to evaluate health in terms of its impact on opportunities.[1] But the most common procedures used to assist medical resource allocation and priority setting decisions do indeed attempt to evaluate health in terms of its impact on well-being. Health states are ordered on the basis of their contribution to the well-being of the people who are in those health states.

Perhaps the best-known procedure to evaluate health states is to use QALYs, or Quality Adjusted Life Years. There are different techniques for the elicitation of the preferences which determine the value of health states in QALYs. The health states are usually described in terms of levels of ability and physical, psychological, and social functioning.[2] Thus, when respondents form their preferences over health states, perhaps the most important factor they take into account is the impact of health: they evaluate health states in terms of their consequences for well-being.

The descriptions of health states in QALY measurement are often constructed from health-related quality of life (HRQOL) instruments. But there are also many HRQOL instruments which do not follow the QALY methodology. Many of them use questionnaires to directly evaluate the impact of specific conditions and treatments on the functioning of patients. The questions focus on those aspects of life which are affected by the condition. In some cases, the respondents are explicitly asked to evaluate their circumstances by indicating their satisfaction with them, given their health status. Often, the questionnaires also include items on adjustment, integration, and social support.[3]

A variant of QALY measurement, although with important methodological differences, is carried out on a large scale. The World Health Organization measures the burden of disease—the harm of mortality and disability from disease and injury for the population of different countries—on the basis of Disability Adjusted Life Years (DALYs). In DALY measurement, health states are described in medical terms. The value of these health states is determined by their "disability weights." The weights were determined by an international panel of experts. When they

[1] This view is developed by Norman Daniels [7] .

[2] See Froberg and Kane [8] and Nord [17]. Some of the preference elicitation techniques are briefly described in note 15.

[3] See Bowling [1].

assigned the weights, the experts were instructed to take into account the social environment in which people with a given condition have to function. The intention of the developers was that DALYs should reflect more than impairment (the manifested disease or pathology), but less than handicap (the overall consequences of a disability in a given social environment). The notion of disability aims to capture something in between: it reflects only the impact of an impairment on well-being. The weights express how well a person's life goes with respect to the given disability only.[4]

All of these approaches attempt to strike a middle ground between measuring the amount of health on the one hand, and measuring overall well-being, determined by factors beyond a person's health, on the other. QALY measurements restrict the states of affairs over which preferences are elicited to health states. Many HRQOL instruments are disease- or treatment-specific. DALYs focus on disability rather than impairment or handicap. Their common background assumption is that it is possible to evaluate health in terms of its impact on well-being by measuring only the health-related part of well-being.

There is, however, a fundamental difficulty with this idea: it assumes that health can be separated from other aspects of well-being. More precisely, it assumes that the value of a person's health for that person can be separated from the value of other components of that person's well-being.

This point has been forcefully argued by John Broome [4]. Suppose that a person's well-being at a time is given by the function

$$w = w(h_1, h_2, \ldots, h_m, d_1, d_2, \ldots, d_n),$$

where $h_1, \ldots, h_m$ are those factors of the person's well-being that constitute her health, and $d_1, \ldots, d_n$ are all the other factors that make up her well-being. If health factors were separable from non-health factors, then this function could take the form

$$w(h(h_1, h_2, \ldots, h_m), d_1, d_2, \ldots, d_n),$$

and $h(h_1, h_2, \ldots, h_m)$ would be the measure of health's contribution to well-being. But health and other aspects of well-being are not separable this way.

Broome gives a simple example. Suppose a person's well-being is the function of the quality of her sight and hearing on the one hand, and the number of books and CDs she has, on the other. The better the person's sight, the more benefit she gets from her books; the more books she has, the more benefit she gets from her good sight. Similarly, the better her hearing, the more benefit she gets from her CDs; and the more CDs she has, the more benefit she gets from her hearing. If h_1 is the person's sight, h_2 her hearing, d_1 the number of her books and d_2 the number of her CDs, this means that her well-being is given by

$$w = w(h_1 d_1, h_2 d_2),$$

[4] See Murray [15].

in which case the contribution of health factors cannot be separated from the contribution of other factors.[5] Because of the inseparability, we cannot measure the value of health in terms of its impact on well-being, since changes in health-related aspects of well-being influence, and are influenced by, other aspects of well-being. Hence the measure does not represent health's contribution to well-being; rather, it is a particular measure of well-being. Therefore, there is no middle ground between measuring the amount of health and measuring overall well-being. The common assumption of standard approaches to evaluate health is problematic.

What are the implications of the inseparability of health and other aspects of well-being for evaluating health? Broome's conclusion is that because of the inseparability problem, we cannot evaluate health as a component of well-being. What we should do instead is to measure the whole of people's well-being—that is, their overall well-being. Dan Brock [3], however, argues that the inseparability can be ignored at least for certain applications and in certain contexts—especially if individual differences in health's impact on well-being can be expected to even out in an aggregate measure. In contrast, Daniel Hausman [12] believes that the inseparability provides an argument for the view that health should not be evaluated in terms of its impact on well-being at all. Even though there is broad agreement on the inseparability of health and other aspects of well-being, there is no agreement about its implications.

Each of the alleged implications points to a different way of coping with the inseparability problem. Perhaps we should shift our focus from measuring the value of health to measuring overall well-being; perhaps we can continue with present practice, since the problem does not have significant practical repercussions; or perhaps we should abandon the way we currently attempt to evaluate health altogether. Is there any other way we can cope with the inseparability? In what follows, I shall attempt to answer this question by considering how we can get from a theory of well-being to the measurement of well-being. Section "Well-Being: Goodmakers and Components" explains the distinction between accounts of the components of well-being and accounts of the goodmakers of these components. The distinction corresponds to different questions we can ask about well-being. Section "Well-Being: Components and Indicators" shows that in order to measure well-being, we need to select indicators for the components. Section "Coping with Inseparabilities" addresses the problem of indicator selection. It argues that we may be able to cope with the inseparability problem if we shift our focus to the selection of indicators which can capture the inseparabilities between components. Section "Evaluting Health States by Preferences" briefly examines the use of preferences as indicators of the value of health states in this context.

[5] For the sake of simplicity, we assume that the well-being function is additive: $w = h_1d_1 + h_2d_2$. Sometimes the same assumption is made in HRQOL measurement when a summary value of quality of life is determined by adding up the scores obtained for different aspects of life. QALYs and DALYs are derived in more complex ways, but this does not affect the argument.

Well-Being: Goodmakers and Components

Health is not the only aspect of well-being which is inseparable from others. Consider enjoyment and accomplishments for example. Other things being equal, the more enjoyment the person derives from her accomplishments, the better life goes for that person; and the more important it is for her to successfully complete a project, the more satisfying and enjoyable that accomplishment is likely to be. The values of accomplishing something and taking enjoyment in that accomplishment are not separable.

It seems that inseparabilities between different components of well-being are pervasive. If so, then the problem goes beyond the relation of health and well-being. This makes it warranted to address the issue from a more general perspective.

So far, I have talked about "aspects" and "components" of well-being as if these notions were unambiguous. But what does it mean that health is a component of well-being? And how can we measure "the whole of well-being," or overall well-being, apart from its components, and *vice versa*? In order to clarify these issues, it will be helpful to distinguish different questions that can be asked about well-being. This is the task of this and the next sections.

When philosophers discuss theories of well-being (or prudential value, to use the technical term), they usually have in mind some theory that explains what it is in virtue of which someone's life goes well for that person (that is, apart from its value for others or from the perspective of some other value).[6] Theories of this sort can be called accounts of *prudential goodmakers*, or accounts of goodmakers, for short. Such accounts explain what it is for something to be prudentially valuable for a person, rather than explaining what actually contributes to the well-being of the person. A prudential goodmaker is that which makes it the case that something is a component of well-being. Perhaps the goodmaker is a feature or property of the component; perhaps it is the relation of the person to the component.[7]

Different theories of well-being can be distinguished by their account of goodmakers. Traditionally, three broad types of theory have been discussed. Many hedonists believe that well-being consists in pleasure; that is, they believe both that only conscious experiences can be good for a person, and only those of these experiences are good for her which are pleasant. This is, very roughly, their account of goodmakers. (Of course, more complex versions of hedonism exist, with more complex accounts of goodmakers.)

On desire or preference satisfaction theories, in contrast, something promotes the well-being of a person if and only if that person desires or prefers that thing. For such theories, what makes something good for a person is that it is the object of a desire or preference. On more complex versions of the theory, the desires or preferences may be restricted to those which are held under some set of appropriate conditions, or they may be those which the person *would* have were she in the appropriate conditions to form desires and preferences. The appropriate conditions

[6] Both the technical term and to a large extent the problem of well-being in contemporary philosophy were introduced by James Griffin [10].

[7] I borrow the notion of goodmakers from Moore [14].

may be that the person is fully informed and adequately rational or that she properly appreciates the nature of the object of her desire or preference. Once again, other variants of the theory are possible.[8]

On a third type of theory, the goodmakers are at least partly independent of the person's preferences or mental states. For these theories, some things are good (or bad) for a person in virtue of some feature of those things themselves. Such accounts are traditionally called objective theories, to mark their difference from preference satisfaction or hedonist views, which are considered subjective. There are many versions of this type of theory. What promotes a person's well-being may be what is worthwhile for human beings to want or to seek, or what contributes to a life which is appropriate for human beings to live, or what is rational to care about, and so on. More often, an objective theory is more eclectic: it offers a list of those things which are constitutive of a person's well-being, each with their own, perhaps mutually unrelated, goodmakers. Such views are known as objective list accounts.

There are well-known problems with each of these views, and even their classification may be debated.[9] What is worth noting, however, is that when philosophers discuss well-being, sometimes their focus is not on goodmakers at all, but on particular *components* or goods which promote well-being. The question of just what these components are is different from the question of what it is that makes something good for a person. Consider a simple version of hedonism as an illustration. On this view, there is only one component of well-being: conscious experience. There is also only one goodmaker: the pleasantness of the experience.

Most philosophers would agree that this theory is implausible. They would argue that plausible theories are pluralistic about the components of well-being.[10] Objective list theories are by their construction pluralistic. They may even allow that the items on the objective list vary from person to person.[11] Other objective theories are usually pluralistic as well. And on desire or preference satisfaction theories, anything can be a component of well-being that the person desires or prefers—or would desire or prefer in the appropriate conditions.

Thus, a complete theory of well-being incorporates both an account of goodmakers and an account of components. The inseparability problem pertains to components, and not goodmakers; and it pertains to those theories which are pluralistic about goods. Naturally, if there was only one component of well-being, there would be no inseparability problem. I shall, however, assume that the correct theory, whatever it is, is pluralistic.

[8] The latter formulation is by Griffin [10: 11]. For other versions of this type of theory, see, for instance, Railton [19] and Brandt [2].

[9] Here I put these problems aside. See, for example, Griffin [10], Scanlon [20], and Sumner [21] for discussions of some of the major controversies.

[10] It may be objected that hedonism can be pluralistic about the components of well-being as well. For instance, hedonists may claim, rather than that conscious experiences are the only good, that all goods must be experienced. This excludes the possibility that a person who has false beliefs about her experiences can benefit from her illusory experiences. (This possibility has seemed to many philosophers a central feature of hedonism, and they have objected to it on this ground.) Hedonists, of course, are free to modify their theory this or some other way.

[11] One criticism of objective list theories is that all they do is to provide a list of components of well-being without an adequate account of their goodmakers. See, for instance, Sumner [21: 45–46].

Well-Being: Components and Indicators

Social scientists working on welfare measurement sometimes start their articles by noting, somewhat exasperatedly, that philosophers have disagreed on theories of well-being at least since the time of Aristotle and there is no end in sight to their disagreements. At the very least, this is an exaggeration. Most of the disagreements concern goodmakers; with respect to components, there is actually broad agreement. Even if there remain disagreements, and even if some of these are important, they are much less sharp than the controversies surrounding accounts of goodmakers.

Consider James Griffin's list of prudential values, his theory of well-being. Griffin argues that well-being consists in several irreducible prudential values: accomplishment, understanding, enjoyment, deep personal relations, and the components of human existence—which include values such as autonomy, basic capabilities that enable one to act, and liberty [10: 67–68]. Many philosophers would agree that these items are components of well-being, even if they give a different account of what it is in virtue of which they are components. For instance, if you are attracted to a preference satisfaction theory, you may agree that these goods are objects of informed and rational preferences.[12]

The broad agreement with respect to components of well-being is further attested by the idea that even if you accept an objective list theory, you are likely to hold that people who properly appreciate the nature of the objects of their desires will desire the items on the list. In the words of Thomas Scanlon:

> Someone who accepts a substantive goods theory, according to which certain goods make a life better, will no doubt also believe that these goods are the objects of informed desire—that they would be desired by people who fully appreciated their nature and the nature of life. [20: 190]

This suggests that social scientists need not be unduly exasperated with the lack of progress in philosophy in the theory of goodmakers. Instead, they can build on the broad agreement that obtains with respect to the theory of components. As it is often pointed out, social scientists and policymakers cannot wait until philosophical problems are sorted out. But the agreement with respect to the components of well-being may be used for the purposes of welfare measurement without problematic commitments in the remaining controversies.

In any case, if we are interested in measuring well-being, our interest in theories of well-being is going to be even more indirect than the foregoing discussion suggests. Few components of well-being are amenable to direct measurement. Most have to be measured with the help of *indicators*. An indicator tells you about people's access to, or possession of, the good it represents. It should be able to capture the changes in the level of people's well-being with respect to the

[12] Griffin argues that if the informed preference (or desire) satisfaction theory is worked out fully, then it becomes indistinguishable from a substantive account of well-being. (See his remarks in Crisp-Hooker [5: 281–285].) He thus believes that familiar distinctions between reason and desire or subjectivity and objectivity are untenable with regard to theories of well-being. Thus, the question whether the items on the list are identified by the proper appreciation of their nature or they are identified by the desires formed on the basis of the proper appreciation of their nature ceases to be salient. See Griffin [11: 32–36].

corresponding good. Ideally, it should be possible to express it in quantitative form and it should be comparable across people and populations.

Consider again Griffin's list. Griffin does not distinguish goodmakers and components: he offers his list as the outcome of deliberation about the ends of life—about what is worth valuing in life—but he also thinks the items on the list are good-making features of life. Neither does he address the problem of indicators, even though most items on his list cannot be measured directly; at the very least, their direct measurement would be infeasible. In some cases, the measurement would be prohibitively costly. For instance, a large-scale survey of people's levels of enjoyment would be very expensive. It would also raise numerous methodological issues from the sincerity of the answers to problems of comparability. In other cases, it would be impossible to carry out direct measurement. What would an adequate methodology for measuring levels of understanding be like? You would probably run into ethical problems as well: people may find your attempt to learn more about their personal relations unacceptably intrusive.

The main problem, however, is simply that these components of well-being are too abstract. One way to overcome this problem is to expand the list. Mozzaffar Qizilbash [18] attempts to operationalize a variant of Griffin's list this way. His own list of prudential values contains the following items:

> (i) minimum levels of health, nutrition, shelter, sanitation, rest and security; (ii) certain basic mental and physical capacities and literacy; (iii) some level of self-respect and aspiration; (iv) enjoyment; (v) autonomy; (vi) liberty; (vii) significant personal relations and some participation in social life; (viii) accomplishment and; (ix) understanding. [18: 2011]

This expanded list of components contains items that people value instrumentally as well as items which are thought to have final value. Griffin's list only contains items that are valued as ends (we discover them through deliberation about the ends of life). Qizilbash includes the extra items on the basis that they are necessary conditions for a life to go well. One argument for this expansion may be that as long as the instrumentally valued items are means to prudential values and they are held universally or generally enough (and perhaps, furthermore, if they are readily amenable to measurement), they should be included in the interests of operationalizability. The distinction between goods that are valued only instrumentally and goods that are valued as ends is unimportant for practical purposes anyway; many components of well-being are actually valued both ways.[13]

Still, many problems remain even after expanding the list of components. The root of these problems is that Qizilbash fails to give an account of the relation of components and indicators. Although he proposes some conventional indicators for

[13] Qizilbash's argument is different. He thinks that since Griffin includes minimum material provision—which has only instrumental value—as a sub-item of the components of human existence, we are free to include other necessary conditions of well-being as well. But actually Griffin mentions having the minimum material goods only as an example of what might be indispensable for having the basic capabilities that enable one to act. He lists it neither as an item nor as a sub-item. So introducing instrumentally valued items on the list is a genuine innovation on the part of Qizilbash. See Qizilbash [18: 2011].

his list, it is unclear what makes something an adequate indicator. For example, for the items under (i), he suggests life expectancy, mortality and adult literacy rates. Arguably, more is involved in these components than what these indicators reflect. In some cases, the relationship of an indicator and what it is supposed to be an indicator of is tenuous: using mean years of schooling as an indicator of understanding or hours of leisure and consumption per capita as indicators of enjoyment are such examples. Qizilbash also concedes that some of the items may not be measurable by any indicator at all. Furthermore, certain items, like literacy, serve both as components and indicators.

It is also peculiar that some items on Griffin's list are expanded, while others are not. Perhaps this is because it is difficult to identify, for example, the necessary conditions for accomplishment or understanding—or they cannot be associated with anything that people universally or generally value as their means. But often it may be possible to identify what *instantiates* these values given the social and economic environment or a person's life conditions. What constitutes accomplishment varies with the opportunities available in particular societies. A person's opportunity for enjoyment depends, among other things, on that person's state of health. Accomplishment-given-one's-opportunities and enjoyment-given-one's-health may be called *complex components* or complex goods. Thus, there is another way in which the array of components may be expanded for practical purposes: it can include complex components for which we can seek indicators. In order to be able to operationalize a list like Griffin's, it seems it must be expanded both ways.

Operationalizability always comes at a cost. If complex components are included among the components of well-being, then the number of components for which we have to find indicators threatens to become unmanageable. It becomes unfeasible to select indicators for all of them; we have to focus on some and exclude others. So some information is bound to be lost in the process of operationalization. But this problem crops up anyway. For instance, if we insisted on finding indicators for the values on Griffin's original list, a lot of information would be lost just as well, since the items are too abstract. As a matter of fact, some information may be gained by selecting indicators for less abstract components, since they might be more relevant for the purposes of a particular study. Expanding the list of components allows us to balance these gains and losses in information content.

I have used Qizilbash's extension of Griffin's list of prudential values to illustrate the way components of well-being are associated with indicators. The theory of goodmakers they both accept does not matter for the operationalizability of their theory of components. An informed desire satisfaction theory, as I argued above, may accept a similar list as the objects of informed desires. These objects of desire have value as ends. In order to measure well-being, other objects of desire have to be taken into account. Some of these will be desired only instrumentally; others will be desired as instantiations of other goods. Thus, the same problems of operationalizability arise if we start from some other theory of well-being.

Coping with Inseparabilities

Broome's argument for the inseparability of health and well-being concerns components of well-being. He is interested in how particular goods determine a person's overall well-being and whether it is possible to measure only part of a person's well-being. As I have argued, however, in order to measure well-being, we have to consider indicators rather than components. Inescapably, this means that we have to address the relation of components and indicators.

One possibility is that the relation is identity. But I have already shown in the last section that this is usually not the case. While we perhaps have direct access to some of the goods which determine well-being, in most of the cases our access is indirect.

A second possibility is that the relation between indicators and components is causal. The idea here would be that even though we cannot directly access some particular component, we can access and measure something that is the effect of the possession of, or the access to, the component we are interested in. Particular goods may have effects on a number of different things that might be used as their indicators. Thus, for example, life expectancy at various ages may be influenced by the components under (i) on Qizilbash's list.

But it is unrealistic to expect that we can identify precise causal relations between particular components and their indicators. The causal direction may go both ways, making it difficult to separate cause and effect. Moreover, many of the indicators used in health evaluation encompass more than the factors which bear directly on people's health status. For instance, life expectancy may be influenced by the crime rate or workplace safety regulations. Actual indicators used in the measurement of well-being go beyond easily identifiable causal relations. This suggests that the relation is looser.

Indeed, on what I shall call the *extreme view*, the question about the relation of indicators and components does not even make sense. In the context of health indicators, this view is proposed by Anthony Culyer [6: 5–9]. Culyer distinguishes two conceptions of health: the traditional medical approach and the "characteristics" approach. The medical approach defines health as the absence of disease or pathology. The characteristics approach describes health in terms of people's characteristics which are selected on the basis of their importance in the given social, economic, and cultural environment.

Culyer argues that if you accept the characteristics approach—which is well on the way to becoming the dominant one—then you have to accept the extreme view:

> The indicator incorporating various characteristics *is* "health" (or "ill-health"). ... [T]he question whether the indicator is a good measure of *health* is meaningless. The appropriate question is whether the right characteristics are included (requiring a set of value judgments in the normative case and a set of empirical judgments in the predictive case). [6: 9, his emphases]

On this view, the only constraints on the selection of indicators are, on the one hand, that they incorporate acceptable moral values and that they give a consistent ordering of health states (the normative case). On the other hand, indicators must also be good predictors of behavior, including behavior which is relevant to policy

decisions—for instance, the utilization of health services (the predictive case). A good indicator is morally acceptable, consistent, and useful for the purposes of decision making.

The extreme view denies that an indicator, in addition to these conditions, must also be *valid*. Validity concerns whether an indicator or measurement tool represents what it is intended to measure. Culyer, as a matter of fact, denies that *content validity* is necessary for indicators for the evaluation of health Culyer [6: 9]. Content validity, roughly, is the correspondence of the indicators to the underlying concept which they are intended to represent.[14] On the medical approach to health, it makes sense to ask whether an indicator really corresponds to some aspect of health. But on the characteristics approach, health is identified with the selected indicators. Consequently, there is no point in inquiring into the relation of these indicators to some underlying concept. There is no correspondence to consider.

But this cannot be right. The characteristics approach does incorporate something health indicators correspond to: people's characteristics which are deemed important, with respect to health, in some social, economic, and cultural environment. The notion of "characteristics" is understood broadly, and it applies both on the individual and social level. Thus, Culyer says, "a one-legged professor is less handicapped than a one-legged footballer" [6: 8]. What he must mean, however, is that being one-legged is worse for a footballer than for the professor, and this is because of the difference in some relevant "characteristic." If this is so, then we must be able to give some *reason* why differences with respect to this characteristic are important enough to merit the selection of an indicator (which, on his view, defines health). Content validity has precisely to do with whether we can give reasons like that. Therefore, the extreme view is untenable. The correct view is more moderate: it holds that the relation of indicators and components is *normative*.

I think the root of the problem with Culyer's view is that he distinguishes two *conceptions* of health, rather than making the distinction between measuring the *amount* of health and measuring the *value* of health.

To see this, note that Culyer's attempt to locate the distinction within conceptions of health, rather than its measurement, is actually quite common. It will be helpful to consider a slightly different attempt. Robert Kaplan [13] draws a similar distinction in terms of the goals of health care. He distinguishes the "biomedical model" and the "outcomes model." The former is concerned with disease and diagnosis. The goal of health care, on this view, is the alleviation of symptoms and the eradication of disease. Health benefits are defined in terms of changes of diagnosis. In contrast, on the outcomes model, the goal of health care is to help people to live longer and better quality lives. Health benefits are defined comprehensively.

[14] In contrast, reliability, which was referred to in the last paragraph, concerns the extent to which the measurement based on some set of indicators produces consistent results. Both are commonly regarded as necessary conditions for sound measurement tools. There are different types of reliability and validity, and different methodologies of establishing them for indicators and measurement procedures. As opposed to other types of validity, establishing content validity is largely a conceptual matter. In the context of health evaluation, validity and reliability are discussed in Froberg and Kane [9].

Kaplan's biomedical model of evaluating health care corresponds to Culyer's medical conception of health: if health is defined as the absence of disease and pathology, then the proper goal of medicine is to alleviate symptoms and eradicate disease. The outcomes model, on the other hand, corresponds to the characteristics approach: if the aim is to make people's lives better with respect to health, then we should be concerned with those characteristics of people that are important in the given social, economic and cultural environment. The assessment of health care, in the former case, requires measuring the amount of health; in the latter case, in contrast, it requires measuring the value of health—most commonly, evaluating health by its impact on well-being.

Thus, it seems to me that the intuition behind the distinctions between conceptions of health and the goals of health care are parasitic upon the distinction between measuring the amount of health and evaluating health. Rather than multiplying conceptions of health, Kaplan's and Culyer's distinctions should be substituted with the distinction between measuring the amount and the value of health. This distinction allows us to recognize both that what is important is the value of health for people, and that health care should be responsive to the impact of health on well-being.

In order to evaluate health in terms of its impact upon well-being, we need to select appropriate indicators. We should, furthermore, take the more moderate view on the relation of indicators and components of well-being; specifically, the question about the content validity of indicators makes sense. Furthermore again, the indicators should correspond to identifiable components or goods—but the list of components should be defined in the broad way I have argued for. It should include goods which are valued instrumentally as well as goods which are instantiations of more general components.

Return to Broome's example. Health contributes to well-being in a number of ways. A person's sight influences the benefit she gets from her books; her hearing influences the benefit she gets from her CDs. Given the inseparabilities, we want to select indicators which can capture these benefits. A more clumsy way of saying this is that we want to select indicators for the complex components of having-some-number-of-books-with-some-level-of-sight and having-some-number-of-CDs-with-some-level-of-hearing.

While perhaps sight and books are, these complex components are certainly not goods that can be directly assessed. We need some indicator for them. Perhaps the person's evaluation of their benefit or her preferences regarding their benefit can serve as indicators. The point I want to make is that if this is so, then the problem of the inseparability of health and other aspects of well-being (as well as other inseparabilities) can be transformed into a problem of the selection of indicators.

As I argued, we have to use indicators to measure well-being. Because of the inseparability problem, it is not sufficient to take different components of well-being independently into account when we select indicators. We also have to consider inseparabilities between components and select indicators which are able to reflect them. When we evaluate health, we have to select indicators (if they exist) which are appropriate for the (complex) components in which health plays a part. This

way, we try to cope with the inseparability problem by shifting it to the selection of indicators.

Of course, it is possible that the inseparabilities between components of well-being resurface in inseparabilities between indicators for these components. But this is not a *conceptual* problem in our account of the components of well-being, but a *measurement* problem for indicator selection—a problem which we may or may not be able to solve (or at least circumvent) with the tools at the disposal of empirical research.

This proposal for coping with the inseparability of health and well-being has several advantages. First, it should not be more difficult to apply in practice than its alternatives. Remember that Broome suggested that because of the inseparability, we have to measure overall well-being. But, in practice, measuring overall well-being is going to be similar to what I have described above. Even if you accept some theory on which there is only one component—a simple form of hedonism, for instance—it is likely that you have to utilize some indicator to be able to measure overall well-being. After all, a person's mental states are not directly accessible. Thus, a simple hedonist would have to rely, for instance, on people's reports of their experiences. More complex theories usually accept that there are many different components. Since most of these are likely to be directly inaccessible, any measurement attempt on the basis of such theories will have to make use of indicators.

Second, the question of what particular indicators are appropriate in the context of the evaluation of health (or in other contexts) is left open. This is as it should be. The choice of indicators should to a large extent be governed by empirical considerations—although, as opposed to what the extreme view suggested, there are also conceptual constraints on the candidate indicators. But most of the questions encountered in indicator selection are empirical: whether the indicators show adequate correlation with other important variables, whether they give consistent results over time and across studies, whether they are feasible to use, and so on. These questions are part of ongoing empirical research programs.

Finally, the proposal gives an explanation for the divergence in opinion regarding the implications of the inseparability of health and other aspects of well-being. If coping with the inseparability problem hinges on the selection of indicators, then the extent and precision to which it is possible to measure the value of health with respect to its impact on well-being is a matter of how successful we are in selecting appropriate indicators. If it turns out that the ability of the best indicators that we can identify to cope with the inseparabilities is limited, then we can evaluate health only approximately and with a lot of information loss. Whether this can be tolerated or ignored (for instance, by arguing that the lost information is not relevant) depends on the context. If it turns out to be impossible to find appropriate indicators, then we might either have to shift the focus of measurement or, as Hausman suggests, give up the idea of evaluating health in terms of its impact on well-being. Indeed, Hausman argues that the most common indicator in health evaluation is unsound. This issue is taken up briefly in the next section.

Evaluating Health States by Preferences

In practice, health states are usually evaluated on the basis of preferences. In order to establish QALY and DALY values for different conditions, respondents are asked to indicate trade-offs between quantity and quality of life.[15] The elicited preferences are likely to be complex: they express the judgments of the respondents on the basis of the information or experience they have about different conditions, their beliefs about the consequences of these conditions in their social, economic, and cultural environment, their feelings about the badness of them, and doubtless many other factors besides. In health-related quality of life surveys, respondents are often asked directly for their judgments about their functioning, levels of pain, mobility, or ability to carry out everyday tasks.

When these judgments and preferences are used, it is assumed that they can serve as adequate indicators for establishing the relative value of different conditions. That is, they reflect the health-related component of well-being adequately. But this assumption can be questioned. It can be pointed out that people may have false beliefs or that they may make mistakes of reasoning when they form their preferences. They may also fall prey to framing effects and other heuristics and biases, or their judgments may be influenced by social expectations.

These problems pertain to the way respondents arrive at their preferences and judgments, and perhaps it can be argued that some of these pitfalls may be avoided by careful research design. But some critics have also pointed out that there may be problems involved not only in the way judgments and preferences are arrived at, but also in their basis. For instance, Daniel Hausman argues that people's preferences over health states may depend on factors other than the impact of those health states on well-being. Many preferences, though perhaps not all, are based on judgments: if challenged, the person who holds these preferences should be able to defend them by citing the relevant reasons. But if all we do is eliciting preferences, how can we make sure that the underlying reasons are good and relevant? And if health states should ultimately be evaluated on the basis of these reasons, why do we need to elicit preferences at all?

So there are at least three problems with evaluating health on the basis of preferences. First, something may go wrong with the formation of preferences. Second, even if the preferences are based only on true beliefs and impeccable with respect to the way they are arrived at, they may still reflect the wrong sorts of reasons. Relying on preferences conceals these underlying reasons. Third, even if the preferences reflect relevant and good reasons, it is unclear why they cannot be

[15] More precisely, they may be asked to form preferences regarding the time they would be willing to sacrifice to live in a healthy condition rather than with diminished health (time trade-off technique). Or they may be asked to consider the alternatives of living in a diminished health state for some number of years on the one hand, and living in full health for the same number of years but with some probability p of death, on the other, and indicate the value for p at which they would be indifferent between the alternatives (standard gamble technique). In DALYs, variants of the person trade-off technique are used: respondents are asked to trade-off life extensions (or health improvements) for some number of people with some condition and some number of people in full health.

sidestepped. Perhaps health evaluation can go directly to the reasons that determine the value of health states without having to worry about preferences.

Hausman's example of "Moneyland" illustrates these problems nicely [12: 265]. Suppose that the denizens of Moneyland evaluate health exclusively in terms of its impact on income. When health economists ask them to compare health states, they base their preferences on their beliefs about the effects of these health states on income. Are these preferences adequate indicators? They may not be if they are based on false beliefs about the impact of health. But suppose these beliefs are true. They may also be inadequate indicators if they reflect the wrong sorts of reasons; but suppose health should indeed be evaluated in terms of its impact on income. The question that remains, however, is why the health economists of Moneyland should elicit these preferences: they can provide the reasons themselves. They can map the impact of health states on income and measure the value of health directly. So the problem is not so much that preferences are inadequate indicators, but that they are redundant as indicators.

This picture of the evaluation of health is very different from the one I have been discussing. If health is to be evaluated in terms of its impact on well-being, then it denies that indicators are needed at all: we can go directly from a theory of well-being to its measurement. Armed with such a theory, we can identify the reasons which determine the value of health states. But, of course, no generally accepted theory of well-being exists. Perhaps we can identify relatively uncontroversial components and attempt to discover how they are supported by reasons. But it is difficult to see how this can be done. There are many relevant components, and different components may be relevant depending on factors ranging from the social, economic, or cultural environment down to the circumstances of individuals. Preferences seem well-suited to serve as indicators, since when people form them, they presumably take these factors into account.

Moreover, Hausman's argument relies on an overly sharp distinction between reasons, judgments, and preferences, suggesting that you can get to the first by bypassing the rest. Preferences may often be unsuitable to help getting us to the relevant reasons, but tossing them out altogether makes the task of health evaluation more, rather than less, difficult. In practice, the gap between eliciting preferences and reasons may be narrowed. In some studies, respondents are asked to justify their preferences, efforts are made to provide them with information and to check the consistency of replies. Often, health evaluation involves a "deliberative" process, including group discussion and argument. In assigning disability weights for DALYs, the preferences of the assessors were elicited repeatedly with different methodologies and the assessors had to revise them to root out inconsistencies.

There is also an epistemic worry about Hausman's proposal. Why should we think that health economists and other experts have better epistemic access than others to the reasons which are relevant to the value of some health state—including those who are in those health states? If we were interested in measuring the amount of health, as opposed to evaluating health, perhaps it would be warranted to rely on their expertise. But if we are interested in evaluating health in terms of its impact on well-being, then experts do not have inherent advantages to judge, for example, the

impact of health on the well-being of people from very different socio-economic backgrounds.

Part of the reason for their lack of epistemic advantage is that health and other aspects of well-being are inseparable. This implies that health's contribution to well-being is not fixed: it cannot be settled once and for all by, say, a panel of competent judges. Thus, the discovery of the reasons which should inform health evaluation is an ongoing exercise.

Leaping into the Swamp

Hausman does not think that health should be evaluated in terms of its impact on well-being at all. One of his reasons is that this approach is bound to be controversial, for there remain numerous disagreements about theories of well-being. He says that

> the view that health should be evaluated by its bearing on well-being runs into trouble when one thinks hard about well-being. If one is looking for solid ground, why leap into a swamp such as the theory of well-being, where doubt and controversy reign? [12: 273]

I have, perhaps unwisely, made that leap. Since the idea that health should be evaluated by its impact on well-being is intuitively attractive, we have to see whether we can swim or sink in that swamp. I have argued that if we distinguish different questions about well-being, we can transform the problem of the inseparability of health and other components of well-being to the problem of the selection of indicators. Thus, the answer to the question of what the implications of the problem are depends on whether we can identify appropriate indicators. How much of a practical problem the inseparability causes depends on these indicators.

Acknowledgement I would like to thank Dan Brock, John Broome, Norman Daniels, Dan Hausman and two anonymous referees for detailed comments. I would also like to thank an audience at the Harvard University Program in Ethics & Health for useful discussion.

References

1. Bowling, A. (2001). *Measuring disease: A review of disease-specific quality of life measurement scales*, 2nd edn. Philadelphia: Open University Press.
2. Brandt, R. B. (1979). *A theory of the good and the right*. Oxford: Clarendon Press.
3. Brock, D. W. (2002). The separability of health and well-being. In Murray et al. [16: 115–120].
4. Broome, J. (2002). Measuring the burden of disease by aggregating well-being. In Murray et al. [16: 91–113].
5. Crisp, R., & Hooker, B. (Eds.). (2000). *Well-being and morality: Essays in honour of James Griffin*. Oxford: Clarendon Press.
6. Culyer, A. J. (1983). Introduction. In A. J. Culyer (Ed.), *Health indicators* (pp. 1–22). Oxford: Martin Robertson.
7. Daniels, N. (1985). *Just Health Care*. Cambridge: Cambridge University Press.
8. Froberg, D. G., & Kane, R. L. (1989). Methodology for measuring health-state preferences—I: Measurement strategies. *Journal of Clinical Epidemiology, 42*, 345–354.

9. Froberg, D. G., & Kane, R. L. (1989). Methodology for measuring health-state preferences—II: Scaling methods. *Journal of Clinical Epidemiology, 42*, 459–471.
10. Griffin, J. (1986). *Well-being: Its meaning, measurement, and moral importance*. Oxford: Clarendon Press.
11. Griffin, J. (1996). *Value judgment: Improving our ethical beliefs*. Oxford: Clarendon Press.
12. Hausman, D. M. (2006). Valuing health. *Philosophy and Public Affairs, 34*, 246–274.
13. Kaplan, R. M. (2003). The significance of quality of life in health care. *Quality of Life Research, 12* (Suppl.), 3–16.
14. Moore A. (2000). Objective human goods. In R. Crisp & B. Hooker (Eds.), *Well-being and morality: Essays in honour of James Griffin* (pp. 75–89), Oxford: Clarendon Press.
15. Murray, C. J. L. (1996). Rethinking DALYs. In C. J. L. Murray & A. D. Lopez (eds.), *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*, (pp. 1–98). Cambridge, MA: Harvard School of Public Health on behalf of the WHO and the World Bank.
16. Murray, C. J. L., Salomon, J. A., Mathers, C. D. & Lopez, A. D. (Eds.). (2002). *Summary measures of population health: Concepts, ethics, measurement, and applications*. Geneva: World Health Organization.
17. Nord, E. (1999). *Cost-value analysis in health care: Making sense out of QALYs*. Cambridge: Cambridge University Press.
18. Qizilbash, M. (1997). Pluralism and well-being indices. *World Development, 25*, 2009–2026.
19. Railton, P. (1986). Facts and values. *Philosophical Topics, 14*, 5–31.
20. Scanlon, T. M. (1993). Value, desire, and the quality of life. In M. C. Nussbaum & A. Sen (Eds.), *The quality of life* (pp. 185–200). Oxford: Clarendon Press.
21. Sumner, L. W. (1996). *Welfare, happiness, and ethics*. Oxford: Clarendon Press.