

Wolff, J; Edwards, S; Richmond, S; Orr, S; Rees, G; 'Evaluating interventions in health: a reconciliatory approach', **Bioethics** (2012), 26:9, pp.455-463. [10.1111/j.1467-8519.2011.01888.x](https://doi.org/10.1111/j.1467-8519.2011.01888.x)

ARTICLE

Evaluating Interventions in Health: A Reconciliatory Approach

Wolff, J; Edwards, S; Richmond, S; Orr, S; Rees, G

UCL (University College London), Gower Street, London, WC1E 6BT
j.wolff@ucl.ac.uk

ABSTRACT

Health-related Quality of Life measures have recently been attacked from two directions, both of which criticise the preference-based method of evaluating health states they typically incorporate. One attack, based on work by Daniel Kahneman and others, argues that 'experience' is a better basis for evaluation. The other, inspired by Amartya Sen, argues that 'capability' should be the guiding concept. In addition, opinion differs as to whether health evaluation measures are best derived from consultations with the general public, with patients, or with health professionals. And there is disagreement about whether these opinions should be solicited individually and aggregated, or derived instead from a process of collective deliberation. These distinctions yield a wide variety of possible approaches, with potentially differing policy implications. We consider some areas of disagreement between some of these approaches. We show that many of the perspectives seem to capture something important, such that it may be a mistake to reject any of them. Instead we suggest that some of the existing 'instruments' designed to measure HR QoLs may in fact successfully already combine these attributes, and with further refinement such instruments may be able to provide a reasonable reconciliation between the perspectives.

1. Introduction

Mainstream health economists measure the seriousness of various (ill-)health states comparatively: members of the general public are asked about which of various health states they would prefer if they were forced to choose between them. The values that emerge from these rankings (often referred to as measurements of 'Health Related Quality of Life' (HRQoL)) are then

used to answer questions of distributing scarce health care resources by incorporating them in calculations of Quality Adjusted Life Years (QALYs).¹

However, some critics, notably Paul Dolan and Daniel Kahneman, have questioned this preference-based methodology. They argue that it is better to focus on people's actual *experiences* of the various health states in question.² A different type of criticism argues for some more objective form of evaluation. A number of authors, inspired by Amartya Sen, argue for a health capability approach, according to which what matters in various health states is how they affect or restrict the individual's capability to function in key respects.³

Alongside the difference in opinion about what constitutes the value of health states, there are different views **[Page 456 →]** about *who* should be asked to evaluate them. The three possibilities most often considered are: (i) the general public (ii) patients (iii) professional experts. Typically, these options match the three approaches set out above: the preference approach argues that the general public are the appropriate consultees, while the experience approach consults (experienced) patients, and the capability approach elicits the impact on capability to functioning of various health conditions from experts (often clinicians). But these pairings are not necessary; one might argue, for example, that the public and patients, rather than experts, should evaluate capabilities.⁴

These two distinctions, 'preferences/experience/capability' and 'public/patient/professional', cut across a third. Each group of consultees comprises a number of individuals, whose assessments

¹ See, for example, Michael F. Drummond, Mark J. Sculpher, George W. Torrance, Bernie J. O'Brien and Greg L. Stoddart, *Methods of Economic Evaluation of Health Care Programmes*, third edition (Oxford: Oxford University Press, 2005).

² Paul Dolan, and Daniel Kahneman, 'Interpretations of utility and their implications for the valuation of health, *The Economics Journal* 118 (2008): 215 – 234.

³ Daniel Hausman's argument that we should focus on the extent to which various health states 'restrict activities' is, broadly speaking, a capability approach (Daniel Hausman, 'Valuing health: a new proposal' *Health Economics* 19 (2010) 280-96. See also Richard Cookson 'QALYs and the capability approach', *Health Economics* 14, (2005): 817-829 and Jennifer Prah Ruger *Health and Social Justice* (Oxford, Oxford University Press, 2010).

may vary. How do we proceed from these individual judgements to a final output? One approach uses forms of aggregation, such as averaging. Another approach favours the use of group deliberation and discussion, with the aim of reaching a consensus view. Thus we must also distinguish between aggregative and deliberative methods of reaching a single judgement. Combining these distinctions generates a large number of different potential approaches. If we ignore variations as well as possible further 'hybrid' proposals, the 3x3x2 classification generates eighteen distinct approaches.

The philosophical foundations for the various approaches on offer are quite different and in theoretical and philosophical terms the best will be the one that, on reflection, we find to have best identified, and argued for, the 'real' determinants of the value of health states. Although this theoretical debate is of great importance and merits lengthy discussion, it is not the immediate concern of this paper. Our purpose, rather, is to consider, by examining some illustrative cases, how far the implications *in terms of practical policy*, differ across the three approaches. In other words, the question that we are considering is: what difference does it make, if any, which approach we adopt?

This enquiry cannot replace the larger theoretical question, but it can inform us about its urgency. For if it turns out that all measures converge in fact on the same (or very similar) judgements, we could continue to use existing methods (based on the preferences of the public) to value health, while perhaps reserving judgement about whether they are justified in themselves or because they happen to coincide in output with the 'correct' method, whatever that may be.

There are advantages in retaining use of the standard current methodology if it is possible. This methodology, as used, for example by the National Institute of Health and Clinical Excellence, attempts to assess a health intervention in terms of the QALYs that it provides. To do this it asks members of the public to express their relative preferences over differing health states. However it is a practical impossibility to gain data on all the many thousands of recognized conditions. Instead

⁴ Daniel Hausman (2010) op cit.

health states are analysed using what is called the EQ-5D instrument. This analyses health states according to five dimensions. These are 'mobility'; 'self-care'; 'usual activities'; 'pain/discomfort'; and 'anxiety/depression'; each dimension divided into three levels of severity. The idea is not that the EQ-5D instrument provides a conceptual analysis of health, but rather if someone is in an adverse health state it will affect them in at least one of these dimensions. Preferences are elicited from members of the public among a sufficient number of health states to allow values to be inferred for different severity levels in different dimensions, which can then be used to calculate the values of health states that have not been investigated directly.

If there is general convergence in results from the different approaches to valuation then the advantages of retaining the current methodology are clear. It is well-established, and easier to use in some respects than the alternatives. For example, it is easier to obtain large amounts of data from the general public than from patients; the public can express preferences over a range of hypothetical conditions, whereas patients with experience need to be found for each condition. And there are clear methods for converting preferences into the type of cardinal measures needed for resource allocation, whereas there is not yet any comparable consensus about how to quantify experience or capability. Accordingly, if there are in practice no, or small, differences between the methods of valuation, there is good reason to remain with what we have.

On the other hand, if we find that there *are* significant differences in outcome when we use the different measures, that has important immediate consequences for priority setting. In this scenario, the theoretical discussion becomes more pressing. It is instructive, therefore, to start with our question of what difference, if any, it makes which measure is used.

Our argument here is that there are significant differences between the measures in some cases, yet nevertheless it is possible to retain the current methodology, albeit in substantially modified form. To show this we first rehearse existing research that shows significant differences between the approaches, and then we consider methods of reconciliation. However we do not attempt to discuss all possible approaches to measurement as indicated above. A full comparison of these

would be a lengthy and laborious undertaking. Moreover, as some of the potential combinations are difficult to motivate, an exhaustive treatment is not worthwhile. In this paper we try to bring out some important differences between some [Page 457 →] of the approaches by examining three pair-wise comparisons that reflect the options most salient in the literature:

- i) public preferences vs patient experience
- ii) public preferences vs professional judgements of capability
- iii) patient experience vs professional judgements of capability

Having done this we will turn to the task of reconciliation.

2. Public preference versus patient experience

Dolan and Kahneman argue that evaluations should be experience-based rather than preference-based on two main fronts. First, they point to decades of research (much by Kahneman and associates), that casts doubt on the reliability and stability of reported preferences. In particular, they emphasise the distorting influence of ‘focusing effects’. Asking someone to express a preference about a particular health condition (whether or not they have it) requires them to focus their attention on something that may in reality bother those people who have the condition for only a small portion of each day. Where this is the case, the respondent is likely to exaggerate its severity in her report. Second, they draw on a range of studies that show that people suffering from some conditions, including paraplegia and even tetraplegia, defy expectations: their reported happiness levels are not substantially lower than those of a control group of people without these conditions. These results suggest, of course, that there are some conditions to which patients are able to ‘adapt’.⁵ Interestingly, these similarities in experience persist whether the patient has acquired the condition since birth or was born with it.⁶ By contrast, members of the public express

⁵ Paul Dolan and Daniel Kahneman (2010) op cit.; Gary L. Albrecht, and Patrick J. Devlieger, ‘The disability paradox: high quality of life against all odds’ *Social Science and Medicine* 48 (1999): 977-988; and K.M. Hall, S.T. Knudsen, J. Wright, S.W. Charlifue, D.E. Graves, and P. Werner, ‘Follow-up study of individuals with high tetraplegia (C1-C4) 14 to 24 years post injury’, *Archives of Physical Medicine and Rehabilitation* 80 (1999): 1507-1513.

⁶ N. Abrantes-Pais Fde, J.K. Friedman, W.R. Lovallo, and E.D. Ross, ‘Psychological or physiological: why

very different views from these patients. In one UK study of more than 1000 people who were asked how they would regard being 'confined to a wheelchair for the rest of your life' around 28% of people said that it would be as bad as death, and a further 24% regarded it as worse than death.⁷

The paraplegia case involves a severe condition, but the public also 'exaggerates' in less severe conditions, such as osteoporosis.⁸ Other studies reveal cases where the difference goes in the other direction; that is, where the experience is worse than the public imagine. These cases may include mental illness, and chronic pain. And some conditions appear to cause 'maladaptation', where the patient experience becomes worse over time, not better.⁹

Dolan and Kahneman use these findings to emphasise the discrepancy between what they call 'decision utility' revealed by expressed preferences (whether of the public or patients), and 'experience utility', elicited by day-reconstruction methods, in which patients log the quality of their experience during periods of the previous day. Two of our sets of distinctions are in play here: Dolan and Kahneman contrast experience-based measures to preference-based measures, and patient consultees to the general public, arguing in each case for the former.¹⁰ In other words, they reject both elements of the currently dominant public preference approach.

Given the significant differences found in the studies to which we have referred, one would expect the policy implications of a decision to move from the current approach to an alternative patient

are tetraplegic patients content? *Neurology* 69 (2007): 261-7.

⁷ M.W. Jones-Lee, M. Hammerton and P.R. Philips, 'The value of safety: results of a national sample survey', *The Economic Journal* 377 (1985): 49-72, p. 54.

⁸ Sherine E. Gabriel, Terry S. Kneeland, L. Joseph Melton, Megan M. Moncur, Bruce Ettinger and Anna N.A. Tosteson 'Health-related quality of life in economic evaluations for osteoporosis. Whose values should we use?' *Medical Decision Making*, 19 (1999): 141-148.

⁹ Relevant studies are summarized in Dolan and Kahneman (2010) op cit.

¹⁰ See also Peter A. Ubel, George Loewenstein, and Christopher Jepson, 'Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public', *Quality of Life Research* 12 (2003): 599-607 and D.G. Fryback, 'Whose quality of life? Or whose decision?', *Quality of Life Research* 12 (2003): 609-10.

experience perspective to be significant. Dolan and Kahneman point out that one consequence would be a shift of resources towards mental health care and pain relief. But given fixed budgets, any 'shift' will entail a reduction elsewhere; presumably, on this approach, in those areas – such as physical disability - where most adaptation is observed. If adaptation is as thoroughgoing as it appears, it is arguable that far less should be spent to help patients overcome or avoid those conditions to which they can be expected to adapt.

Suppose for example that a patient needs a life-saving operation. The surgeon can either perform a cheap procedure that leaves the patient with severe mobility problems, or a more expensive operation from which the patient can be expected to make a full physical recovery. Given the public's aversion to mobility restrictions, the public preference approach is likely to support the expensive operation, while reports of (disabled) patient experience would shift the balance of opinion towards the cheaper operation. Hence the switch to experience measures would have significant effects (not all obviously welcome). As Dolan and Kahneman observe, 'all else equal, if adaptation to one condition is more complete **[Page 458 →]** than to another, then the former will be given less priority than the latter'.¹¹ Of course the 'all else equal' clause is vital. It may be expensive to provide the circumstances necessary for successful 'adaptation', and the transitional suffering often experienced by patients before adaptation is achieved must also be taken into account.

There are other types of case, not involving adaptation, where we might also expect similar divergence. With some kinds of mental dysfunction (including some Alzheimer's cases), the patient's experience is not unhappy, yet members of the public typically express a strong preference against the condition.

In light of these cases it is worth thinking about the reasons for some of some of the public's

¹¹ Dolan and Kahneman (2010) op cit. p. 216.

preferences: do they 'exaggerate' the badness of some experiences, or are they subscribers to the 'capability' approach, preferring a life with fewer restrictions even if it does not guarantee a better subjective experience?¹² We return to these issues below.

3. Public preferences versus professional judgement of capability

Within the philosophical literature capability measures were introduced partly in response to the very issue of adaptive preference.¹³ The 'happy slave' example shows why subjective reports of well-being or the expression of someone's preferences might both be inadequate guides to value: even if a slave is happy, and prefers slavery to freedom, nevertheless, it is argued, he or she has real interests that are undermined by slavery. The notion of 'capability' puts flesh on the bones of 'real interest'; capabilities are, in Sen's formulation, freedoms to 'do' certain valuable activities, or to 'be' a certain sort of person. Restricting one's range of activities diminishes one's capability. On such a view a disabled person is in one way like the slave, insofar as both are restricted in their activities.

The issue of adaptation is not however directly germane to the comparison between public preferences and professional judgments of capability considered in the present section: it is patients who adapt, and our present comparison leaves the patient's perspective out. For the purposes of this comparison, the question to consider is what differences one might expect between the public's preferences concerning health states and capability-based 'expert' judgments

¹² Some data even suggests that accident victims are as 'happy' as lottery winners. P. Brickman, D. Coates, and R. Janoff-Bulman, 'Lottery winners and accident victims: is happiness relative?' *Journal of Personality and Social Psychology* 36 (1978): 917-927. It is also possible that people may be genetically disposed towards a personal happiness 'reference level', which supports the idea that people adapt to major life events. See A.E. Clark, E. Diener, Y. Georgellis, & R.E. Lucas, 'Lags and leads in life satisfaction: A test of the baseline hypothesis.' *Economic Journal*, 118, (2008). F222-243, and Alexander Weiss, Timothy C Bates, and Michelle Luciano. 'Happiness is a personal(ity) thing: the genetics of personality and well-being in a representative sample', *Psychological Science* 19 (2008): 205-210. See also G. Livingston, C. Cooper, J. Woods, A. Milne and C. Katona, 'Successful ageing in adversity: the LASERAD longitudinal study', *Journal of Neurology, Neurosurgery and Psychiatry* 79 (2008): 641-645. The idea that patient and public perspectives can diverge has been contested by M.G.T. Dolders, M.P.A. Zeegers, W. Groot., and A. Ament, 'A meta-analysis demonstrates no significant differences between patient and population preferences,' *Journal of Clinical Epidemiology* 59 (2006): 653-664, although it is not clear that the results of the paper really do demonstrate what is claimed.

about the value of those states.¹⁴ As before, if there is no reason to think that there will be significant differences, then it does not matter for the purposes of policy which approach is used.

Daniel Hausman provides examples to illustrate possible divergences between public and expert judgment: 'Members of some social group might cherish impairments such as epilepsy as signs of divine favor. Their preferences would not be a good guide to their health'.¹⁵ Sceptics might dismiss this case as a typical philosopher's fanciful example, with no bearing on actual policy. Yet there are similar real cases. It is easy to imagine that in a society where female circumcision or foot binding were prevalent, the public might rate the harm of these procedures (even in terms of health) lower than doctors. It could turn out that on a public preference approach, these procedures would have to be regarded as *improving* health in that society. From a more 'objective' standpoint, this result is unfortunate.

Hausman argues that a significant problem with preference-based methodology is that it solicits opinion about matters that should be investigated directly. 'A preference ranking of health states depends on judgments concerning the character and consequences of health states, and it is as liable to be mistaken as the judgments upon which it depends. The evaluation of health states by an individual is a cognitively demanding task, to which preferences provide no shortcut'.¹⁶

Hausman's [Page 459 →] conclusion from this argument is not that we should, instead, rely on professionals, but rather that public consultation should be based on collective deliberation about the functional limitations and impact on experience of health conditions, rather than on the

¹³ Amartya K Sen, 'Utilitarianism and welfarism', *The Journal of Philosophy*, 76 (1979): 463-489.

¹⁴ Note that the DALY measure (Disability Adjusted Life Year), which was first introduced in the *World Development Report* 1993 to assist in the calculation of the 'global burden of disease', while structurally similar to the QALY in its quantified approach, is constructed out of valuations provided by the expert judgments of health workers, rather than the public. The most discussed aspect of the difference between the QALY and DALY approaches is that the DALY approach gives a weighting favouring years of young adult life over years of childhood or elderly life, whereas the QALY approach is neutral in this respect (for discussion, see Franco Sassi, 'Calculating QALYs, comparing QALY and DALY calculations', *Health Policy and Planning* 21 (2006): 402-408). This age-weighting issue is independent of the question of the source of the valuation of health states, which is our focus here. We return to how DALYs are generated later in the paper.

¹⁵ Hausman (2010) op cit., p. 265.

aggregation of individuals' preferences. In the terms of this paper, Hausman's view can, in part, be described as advocating public and patient deliberation on capabilities. Yet in a society in which foot binding is widespread, for example, it is not obvious that public joint deliberation on capabilities will fare any better than individualistic preference elicitation.

Of course, we should not suppose that professionals will, in every case, transcend their cultural circumstances to deliver an 'objective' scientific judgement. A Catholic doctor might refuse to perform an abortion or a Jewish doctor overstate the health benefits of male circumcision. Nonetheless in general one would expect some divergence between public and expert judgment on some health-related matters, whether the public is deliberating individually or together, or expressing its preferences or its assessment of capabilities.

Aside from the admittedly important 'cultural' cases such as foot binding and female circumcision, some studies suggest doctors tend to give lower values than the public in ranking some diseases.¹⁷ Other studies show differences in both directions. One pilot study, for example, found that doctors regarded schizophrenia less seriously than the general public, while at the same time placing greater weight on the adverse side-effects of treatment.¹⁸

It is plausible to explain these discrepancies in terms of the difference in expertise of the consultees: as Hausman says, the judgment of health states is a cognitively demanding task, and members of the public are not especially informed about these states. Is the solution, then, to provide the public with all relevant details about various health states before they are asked to express their preferences? Once they are fully informed, one might think, their preferences will

¹⁶ Hausman *ibid* p. 272.

¹⁷ C.J.L Murray, and A.D. Lopez, 'The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020', *Global Burden of Disease and Injury Series Volume 1* (Cambridge MA, Harvard School of Public Health/WHO/World Bank, 1996), p.30.

¹⁸ Tina T. Lee, Jennifer K. Ziegler, Roger Sommi, Catherine Sugar, Ramy Mahmoud and Leslie A. Lenert, 'Comparison of preferences for health outcomes in schizophrenia among stakeholder groups', *Journal of Psychiatric Research*, 34 (2000): 201-210.

better match the judgements of professionals. But this may be too optimistic: we cannot assume that members of the public will absorb information easily, or be disposed to believe it if it goes against their long held beliefs, or to use it judiciously even if they accept it. Indeed some data suggest that, surprisingly, when the public is 'informed', its estimates of utility fall even further below those of their doctors.¹⁹

4. Patient experience versus professional judgment of capability

In this section we explore the possible divergence between doctor and patient perspectives on health. This is a familiar phenomenon in everyday life, and formal studies confirm it. For example, Jachuk et al. compared assessments of the quality of life of patients undergoing antihypertensive treatment from three groups of people: the patients themselves, their companions, and their doctors. The patients' rating was lower than the doctors and the authors, rather dismissively, suggest that one reason for this may be the patients' 'lack of insight into their own problems'.²⁰

As one might predict, where differences are found between doctors' and patients' views of various health states and medical interventions, they suggest that patients are more concerned with the illness experience, while doctors pay more attention to physiological impairment. In the treatment of arthritis, for example, one study showed that doctors thought joint replacements were the closest thing to a 'cure', while patients were less enthusiastic and preferred to try other less invasive treatments.²¹ Furthermore, over time doctors will see certain aspects of health conditions as 'normal' within a patient population since they see them all the time. This sometimes specifically includes pain to which they become desensitized.²²

¹⁹ Murry and Lopez, (1996) op. cit., p.30.

²⁰ S.J. Jachuck, H. Brierley, S. Jachuck, S and P. Willcox, 'The effect of hypertensive drugs on the quality of life', *Journal of the Royal College of General Practitioners*, 32 (1982): 103–105.

²¹ B.L. Lambert, D.N. Butin, D. Moran, S.Z. Zhao, B.C. Carr, C. Chen, and F.J. Kizis, 'Arthritis care: comparison of physicians' and patients' views', *Seminars in Arthritis and Rheumatism*, 30 (2000): 100-10.

²² Ruth Richardson, 'A necessary inhumanity?' *Medical Humanities* 26 (2000):104-106.

The growing body of 'illness literature' lends qualitative support to this divergence. Arthur Frank recounts his experiences of heart attack and cancer, and laments the narrow focus of most physicians.²³ This raises the difficult question, of course, of precisely what an experience-based approach should include; but even with quite a restricted definition, it is clear from numerous patient's accounts that their sense of what matters is frequently quite different from their doctor's.

As we saw earlier, the divergence between the experience and capability perspectives shows up in cases of disability (such as blindness, deafness or paraplegia). These obviously limit capability, but they do not always affect patients' reports of their happiness. The choice of one of these approaches rather than the other in these cases makes a significant difference to the evaluation of the relevant medical interventions. Another type of case where divergence will show up is where subjective [Page 460 →] experience is poor, even though the patient's other functionings are not impaired: clinical depression and chronic fatigue are examples of this type.

5. Are Any of the Perspectives Dispensable?

We have examined some cases where the perspectives of the general public, patients and professionals differ, and it is also clear that judgments based on preferences, experiences and capabilities will sometimes diverge. The next question to consider is whether any of these perspectives can be dismissed. Is each of them strictly relevant to the formulation of health policy?

It would seem perverse to think that policy makers could dismiss as irrelevant either the perspective of patients, or that of professionals (especially doctors). Patients indisputably possess some special knowledge: of how things feel, and how a health condition impacts on life. Doctors also have special knowledge: of clinical evidence, and of how the body works. Health states are not just a matter of present experience; in many cases they also determine the future. Just as an

²³ Arthur Frank, *At the Will of the Body: Reflections on Illness* (Boston: Mariner Books, 2002).

engineer can point to risks and stresses within a system, which may not be having any present impact, so too can doctors. Medical expertise includes the range of future health possibilities and probabilities, such as whether a condition may deteriorate or create new needs in the future.

If the perspectives of patient and doctor are both valuable for their particular insights, it is less easy to see why this should also be true of the general public. Why should we think the public perspective important, if it lacks both experience and knowledge? It is ironic that current measures rely almost exclusively on eliciting the preferences of the public, when intuitively it seems to be the least authoritative group. There are at least two important questions here. First, does the public have any special insight relevant to the evaluation of health interventions? Second, is there any reason to take their views seriously, even if they lack special insight?

To start with the first question, there is no doubt that the public understanding of what it is like to live with a condition is often poor. Where the divergence between the public's and the patient's view can be explained by such ignorance, there is good reason to dismiss the public view. In the cases of mental illness and chronic pain, for example, the patient perspective, whether experience or preference based, tends to rate these conditions as more severe than the public preference approach. The public perspective can be assumed to be 'ignorant' here, as thankfully only a minority of people experience serious mental illness, or chronic pain (indeed it is often claimed that human beings have a poor memory for pain). In this case there seems no reason to retain the public perspective; indeed, one would expect a member of public to 'correct' her view in the light of further knowledge about the experience of these conditions.

Yet, in other cases it is plausible to think disagreement can survive the acquisition of more knowledge about patient experience. One might become aware, for example, of the success with which many patients adapt to restrictions on their mobility, without this affecting her own strong preference not to lose her own mobility. Such concern can be psychologically robust, even in the

face of full information.²⁴

This divergence might reflect the public's view that, for some health conditions, patient experience does not capture everything important. From an external point of view one can have reason to value some conditions above others even if they are not correlated with greater subjective well-being. As in the 'happy slave' case, the existence of choice, freedom and opportunity seems to matter. Physical immobility reduces one's set of options and even if this still leaves enough scope to lead as satisfying a life as one might have done otherwise, people will in advance usually prefer not to have their options reduced. One's range of choice seems *itself* to be of value and, furthermore, individuals know that they have chosen their present options in the light of the constraints obtaining at that time. Even if they know that they could be satisfied with another (narrower) set of options, from their present point of view that set will appear second best. However 'complete' an adapted life may be in terms of experience, it may still be second best from the point of view of preference.

Thus an awareness of patients' adaptive abilities, and a degree of scepticism about the patient's perspective both appear to be warranted: the public preference view has a claim to be retained insofar as it expresses something important that the 'adapted' perspective overlooks.

Another factor often cited in the explanation of the divergence between public and patient perspectives is fear. As Dolan and White point out, public health is not only concerned with health states, but also with anxiety *about* health states (which also affects other aspects of individual well-being, and shows up in such things as consumption patterns).²⁵ The studies reporting public preference against physical immobility relative to other conditions do not seek to explain the results, but we might speculate that the prospect of major disability arouses great fear. For some people the fear of illness **[Page 461 →]** goes well beyond concern about the intrinsic nature of the

²⁴ Peter Ubel, et al (2003), op cit.

²⁵ Paul Dolan and Matthew White, 'How can measures of subjective well-being be used to inform public policy?', *Perspectives on Psychological Science* 2 (2007): 71-85.

condition, and includes concerns such as the response of other people.²⁶

One response to the impact of prospective anxiety is to point out that it is not itself part of the health state that policy makers are attempting to measure: 'accounting for fear is quite separate from accounting for the real losses from a given health state'.²⁷ But it does not necessarily follow that one should discount it; after all, fear/anxiety is a dimension of health (and included on the EQ-5D scale) and, especially in its consequences on health-related behaviour, it can become relevant to health policy. It is worth noting too that, from the start, an argument used in favour of establishing the UK NHS was that it would reduce the anxiety attaching to illness, and especially people's associated 'money worries'.²⁸

Finally, we need to consider whether, even if the public perspective does not provide any special insight, nevertheless, there might be reason to give it weight. A point often raised is that in a tax or insurance system of health, the people paying for health care are ordinary citizens, many of whom do not fall seriously ill. It might be claimed that those who pay have a right to determine how their money is spent.²⁹ Yet, as Nord argues, it seems wrong for policy makers to respect the views of the general population where they are based on beliefs that are demonstrably false.³⁰ Arguably, the political concern about how the public likes to see its money spent should not impinge on the empirical question of how best to measure the benefits of health interventions.

²⁶ Not all of these concerns are 'irrational'. Sarah Anderson reports that one of the hardest aspects of losing an arm is coping with other people's attitudes towards her disability. Sarah Anderson, *Halfway to Venus* (London: Umbrella Books, 2008)

²⁷ Dolan and White (2007) op cit. p. 372.

²⁸ Aneurin Bevan, *In Place Of Fear* (London: Macgibbon & Kee, 1961).

²⁹ The Panel on Cost Effectiveness in Health and Medicine, convened by the Department of Health and Human Services to improve standardisation of cost-effectiveness methods in health care, recommended the resource decisions reflect the costs and effects of interventions to all members of society and concluded that policy makers should approach a representative sample of informed members of the community. M.R. Gold., J.E. Siegel, L.B. Russell, and M.C. Weinstein, *Cost Effectiveness in Health and Medicine*, (New York: Oxford University Press, 1966), p. 425.

³⁰ Eric Nord, *Cost-Value Analysis in Health Care*, (Cambridge: Cambridge University Press, 1999) p. 88.

5. Reconciling these approaches?

It seems that there is something to be said in favour of each of public preference, patient experience and professional assessment of capabilities (as well as other perspectives we have touched on, such as public deliberation about capabilities). How, then, should we respond in specific cases of divergence? Each perspective has advantages and disadvantages; if none can be ignored, yet - for the same reason - none can be chosen as the only approach to evaluation, the obvious thought is that we should try to include all of them.

Despite the differences between the approaches, there is some promising common ground. To see this, let us assume, for example, that one is persuaded of the capability approach and attempts to think through how it may be implemented. In the standard philosophical presentation of the capability approach it is, of course, necessary to make a list of relevant capabilities. For most philosophers 'health' is itself merely a single category of capabilities, among many others, such as affiliation, control over the environment, and emotional well-being, which form part of a list that is either open-ended, as in Sen³¹ or determinate, as in Nussbaum.³² Hence a health capability approach will need a more fine-grained list of health related capabilities. Furthermore, it is typically part of the philosophical understanding of the capability approach that the different capabilities are, in some respects, incommensurable, at least in the sense that lack of one capability cannot always be compensated for by more of another. If this feature is retained then the capability approach immediately disqualifies itself as a possible basis for assisting decision making between scarce resources in health care, insofar as such decisions require quantitative comparisons. Of course, that could be a reason for a very radical change of methodology, but as long as we remain within existing frameworks what will be needed is not so much an application of the capability approach as a modification.

³¹ Amartya K. Sen, *Inequality Re-Examined* (Oxford: Oxford University Press, 1992), *Development as Freedom* (Oxford: Oxford University Press, 1999).

³² Martha Nussbaum, *Women and Human Development* (Cambridge: Cambridge University Press, 2000).

How might that modification go? First a particular list of ‘health capabilities’ will be needed, and second, the elements on that list will have to be given values. But, in fact, the standard methodology – the EQ-5D scheme – does exactly this. As we noted that scheme analyses health into three levels of severity along the five dimensions of health: ‘mobility’; ‘self-care’; ‘usual activities’; ‘pain/discomfort’; and ‘anxiety/depression’. While the primary motivation for analyzing health into five dimensions is not an explicit endorsement of the capability approach but for pragmatic or instrumental reasons (indeed devices like EQ-5D are referred to as ‘instruments’), nevertheless, if we are looking for a health-related capability account that can actually be used to inform resource allocation decisions, something like the EQ-5D scale, even though it is some distance from a pure capability measure, is not a bad starting point.³³

[Page 462 →]

The next step, however, is to consider how those points on the matrix are to be valued. Although the standard methodology currently bases valuation on public preferences, the capability methodology typically (although not exclusively) consults professionals, to evaluate health states. The difference between the professional and public perspectives, we saw, can depend on a number of factors, such as the level of professional knowledge, as well as differing judgments about what is important about health. But it appears that the capability approach can use the health dimensions set out in EQ-5D as the basis for creating its own ‘instrument’ where valuations of dimensions are supplied by doctors.

And, indeed, we can also see that not only can the public preference, and professional capability perspectives use the EQ-5D instrument, so can the experience perspective. The five dimensions of EQ-5D are all either aspects of a patient’s experience, or will affect their experience, and so a patient can give valuations to each dimension, by considering what trade-offs between different experiences he or she would be prepared to make. It seems, on the evidence considered so far,

³³ Cookson (2005) op cit. It should be noted that many health capability theorists would not accept this proposal, seeing the capability approach as incompatible with straightforward quantitative assessment, and therefore resisting this simplifying move. See for example Jennifer Prah Ruger, (2010) op cit.

that patients will consider the dimensions of 'mobility' and 'usual activities' less serious relative to their valuation of 'pain/discomfort' and 'anxiety/depression' than members of the public and at least some professionals. (It is unclear how 'self-care' would be rated.)

In sum, there can be broad agreement about analyzing health states in terms of the EQ-5D instrument, whether one starts from a commitment to preferences, experience or capabilities. However, when it comes to putting weights on the different dimensions, it seems likely that the different perspectives of public, patient and professional will generate different judgments, at least sometimes. In practice, then, we may find ourselves with cases where there is convergence between all three groups' approaches to valuation, cases where two converge and one is out of step, and - theoretically at least - cases where the three groups give three different judgments. So, while we cannot dispense with any of the three perspectives entirely, we do have a way of drawing them into a single framework, within which we can theorize the differences.

It looks then as if the different approaches to measurement can agree that the EQ-5D captures the important dimensions of health, but will differ in their final judgements about the value of different health states. And, as we have noted, this disagreement has serious consequences for priority setting. How, then, do we reduce disagreement?

In principle, we can divide cases into those where one (or more) perspective is mistaken in some way, and deserves to be removed or modified, and those where each are valid and reflect an understandable difference of assessment. So, for example, it may be that patients sometimes 'over-adapt' to their condition, and at least some adjustment should be made to their valuations. Equally, professionals and members of the public may lack experience of the phenomenology of a condition, and as a result become over-optimistic in their assessments. In other cases patient experience and professional evaluation may simply emphasise different aspects of a condition in different ways. In practice, therefore, we are left with two questions. First, how can we decide whether a divergent perspective is what we can call legitimate (in the sense of not being based on faulty reasoning), and second, how do we deal with divergent legitimate perspectives?

To the first question it would be tempting to try to seek an epistemological principle that would rule out illegitimate assessments. However it is unlikely that any such principle could be found, or, at least, command general assent. Nevertheless there will be what we might think of as 'warning signs': the possibility that a perspective may be based on over-adaptation or ignorance. How, then, do we move from here? To avoid arbitrarily privileging of one group over another, or simple averaging, which has no apparent theoretical justification, we propose, in effect, bringing representatives of all three groups together and shutting them in a room to deliberate until a consensus is somehow reached.

Indeed this is roughly what the DALY (disability adjusted life year) methodology already does, or at least simulates.³⁴ In calibrating values for the DALY WHO convened an 'expert panel' comprising health care providers. However, they did so not because they could give more 'objective' values of capability (although they do try to come up with such a set) but to save time educating others such as patients and the public on how to do the trade-offs. They thought that an educated public could have derived a similar set of valuations just as well and they were especially keen to establish a set of common values which could be used across different cultural settings. The healthcare providers were also explicitly instructed to take different perspectives (including the patient's and the public's) and to reconcile any differences before giving an answer just in order to reach a consensus on common values which is presumed to be an objective indicator of 'social preferences' or what the average individual would prefer. This may combine and so hide all sorts of differences between experience, preference and capability measures, and could result in an ad hoc and unclear combination of perspectives, or alternatively, we hope, provide a model of a type of deliberative process.

[Page 463 →]

³⁴ See footnote 14. Indeed Murray and Acharya themselves point out the similarities between the DALY and capability approach, C.J. Murray and A.K. Acharya, 'Understanding DALYs', *Journal of Health Economics* 16 (1997), pp. 703-730, 723-4.

While the DALY is too crude, in making too few distinctions, to use in health priority setting (as distinct from estimating the global burden of disease) its deliberative basis is very useful. Our modification, however, is to reserve deliberation for a second stage of assessment. That is, in a first stage the EQ-5D would be used to elicit separate QALY values from each of the public, patients and professionals according to variants of the standard methodology. If the resulting values are close, then no further work is needed. If, however, there are significant divergences, we recommend that a deliberative meeting is convened in which representatives of the three groups discuss and attempt to defend their reasons for valuing the states as they do. It is possible – this is an experimental question on which research is needed – that such deliberation can reduce disagreement. If it reduces it to manageable bounds then the problem is solved.

But suppose significant disagreement remains. This could either be because ignorance or over-adaptation is so deep that it cannot be shifted, or because the different perspectives are each legitimate. In practice we cannot distinguish these cases if deliberation does not reduce disagreement. What should then be done? Our tentative proposal is that in the case of disagreement the lowest valuation should be used; i.e. the one that rates the health condition as most serious. A justification for this, at first sight, is ‘consequence of error’. It is better to fund a treatment that should not be funded than to fail to fund a treatment that should be.

However, the picture is more complex. For if we are making allocations from a fixed budget then any decision to fund a contested treatment will draw resources away from elsewhere: in effect from treatments that would have been just on the threshold of cost-effectiveness if the new treatment had not been funded. If there are many contested treatments then, in effect, this could have a serious impact. However, this is always a problem when a decision is made to fund a new treatment: the money for that treatment must come from somewhere. Nevertheless, we feel that our proposal is the least imperfect way of dealing with disagreement. And it remains an empirical question how much disagreement will in fact remain after deliberation.

6. Conclusion

We accept that there are good reasons to think that different approaches to the evaluation of health states can lead to different valuations, which sets a problem of which values to use for priority setting. Our approach has been, first, to propose that such disagreement can be modeled using the EQ-5D scale (or something like it) and second, that where disagreement exists it should be reduced, where possible, by deliberative methods. Where disagreement persists even after deliberation we suggest that policy makers give priority to the approach that considers the condition the most serious. Although not perfect, we feel that this is less imperfect than other possibilities. Further research is needed, however, on how to organize deliberation to best effect, and on its likely outcomes.³⁵

³⁵ We would like to thank xxx.