Lessons Learned by (from?) an Economist Working in Medical Decision Making

Peter P. Wakker, PhD

This article is a personal account of the author's experiences as an economist working in medical decision making. He discusses the differences between economic decision theory and medical decision making and gives examples of the mutual benefits resulting from interactions. In particular, he discusses the pros and cons of different methods for measuring quality of life (or, as economists would call it, utility), including the standard gamble, the time tradeoff, and the healthy-years equivalent methods.
Key words: *risk; time tradeoff; healthy-years equivalent; interdisciplinary work; cost-effectiveness; decision analysis.* ***(Med Decis Making 2008;28:690–698)***

1. INTRODUCTION

This article gives a personal account of my experiences as an economist working in medical decision making. Sections 2 and 3, the first part of the article, describe the origin of medical decision making (MDM) in the early 1980s and my background in economic decision theory (EDT) before 1992, the year I entered MDM. Sections 4 to 6, the second part of the article, describe 3 surprises that were awaiting me when I entered MDM. Sections 7 to 10, the third part, present prospect theory and then describe lessons that can be learned from the aforementioned surprises. The article ends with some speculations about future interactions between EDT and MDM and the improvements that can result for the measurement of quality of life (sections 11 and 12).

2. BACKGROUND OF MDM

MDM took off around 1980 with a paper by McNeil and others.[1] These authors showed how the evaluation of medical treatments—up to then often based on the 5-year survival criterion—can be improved using tools from EDT. These tools were based on expected utility, the prevailing theory in those days. Thus, MDM was based on the ideas of EDT in the 1980s. These ideas were laid down, for instance, in the classic book by Keeney and Raiffa.[2] This book used expected utility as the normative basis for rational decisions, a normative basis that I have subscribed to throughout my life. Krischer[3] surveyed MDM until 1980 from the perspective of EDT.

Up to the 1980s, the best approximation of a descriptive theory was taken to be the normative theory in EDT because deviations from rationality were thought to be erratic and chaotic and not suited for modeling in any sense.[4(p406)] Hence, expected utility was also used for descriptive purposes. It continues to be so in MDM today. Keeney and Raiffa's work has been particularly influential for the widely used Health Utility Index, a sophisticated method for measuring quality of life by aggregating different components that are allowed to interact.[5] This method is based entirely on expected utility theory. Many of the ideas in EDT from the 1980s have since changed in this field, yet can still be found in their original form in MDM today.

A strategy typical of MDM is that the concepts and ideas taken from EDT are used as tools or machines without much consideration of the theory inside. MDM'ers will just try these machines out, see how they work, and if their results are satisfactory then MDM'ers will continue to use them. I will call this approach the black-box approach.

I first encountered the black-box approach in a conversation with Lia Verhoef in 1989. At the

Received 15 November 2007 from the Department of Economics, Erasmus University, Rotterdam, The Netherlands. This article is the written text of a lecture presented at the 29th annual meeting of the Society for Medical Decision Making, 24 October 2007, Pittsburgh, Pennsylvania, for the career achievement award of the Society for Medical Decision Making. The author acknowledges Denny Fryback, Harold Sox, and 2 anonymous referees who made helpful comments. Revision accepted for publication 26 May 2008.

Address correspondence to Peter P. Wakker, PhD, Department of Economics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands; e-mail: Wakker@few.eur.nl.

DOI: 10.1177/0272989X08323916

$$\exists\, U : \mathbb{R} \to \mathbb{R},\ \exists\, W : 2^S \to [0,1] \text{ s.t.}$$

$$(E_1{:}x_1, \ldots, E_n{:}x_n) \mapsto \sum_{j=1}^{n}\big(W(E_{\rho(1)} \cup \cdots \cup E_{\rho(j)}) - W(E_{\rho(1)} \cup \cdots \cup E_{\rho(j-1)})\big)U(x_{\rho(j)})$$

$$\text{(with } \rho \text{ a permutation such that } x_{\rho(1)} \geq \cdots \geq x_{\rho(n)}\text{) represents preferences.}$$

Figure 1 Rank-dependent utility (the "inside of a machine").

time, I was working in the psychology department of the University of Nijmegen in the Netherlands, with Arne Maas, Thom Bezembinder, and others. We would sometimes meet with Lia, Lukas Stalpers, Willem van Daal, and other people from radiotherapy who were interested in decision theory. This was before I had become involved in MDM myself. Lia used a formula for evaluation purposes (the time tradeoff method for measuring quality of life that will be explained later, in equation (4.2)), and I immediately noticed that her formula was not in agreement with any EDT formula that I could think of. I wrote down a formula that was in agreement with EDT (equation (4.3)) and handed it to Lia, telling her that this was a correction of her mistake and trusting that she would use the corrected formula for ever after.

However, a surprise was awaiting me. Lia replied that she would not use my corrected formula. Our personal relationship was good enough, fortunately, to bridge the professional distance that arose between us at that moment. Lia explained her case. She told me that the formula she used was just a machine to her that she had found to work well. What I proposed was just another machine that she did not know and that she would not use until it had been found to work well. I was flabbergasted. What Lia told me that day is something that cannot be learned from literature or conversations but only from experience. I did not know then that a future in MDM was awaiting me and that it would take me years of work before I would be able to experience the truth in Lia's words. Instead, I simply did not understand. "They" use a wrong formula, know it is wrong, are made aware of a corrected formula, but still refuse to use the corrected formula. It was the first time that I had been exposed to such nonmathematical ways of reasoning. Unaware of insights yet to come, I concluded that a mathematical mistake is a mathematical mistake.

EDT'ers primarily work on theoretical models, changing and fine-tuning the parts of their machines to always fit together better. They live inside their machines (*coherence*[6]). MDM'ers take the machines produced by EDT'ers as a whole and try them out in practice. Their black-box approach only judges machines from the outside and does not try to see what is inside (*correspondence*[6]).

3. MY BACKGROUND IN 1992

I graduated in mathematics and was awarded a PhD in economics. My research primarily concerned rank-dependent utility and prospect theory.[7–10] Figure 1 reproduces some formulas of these theories. These formulas need not be read and only serve to give an impression of the world I lived in. The figure shows the insides of a machine, so to speak, and I trust that the medical readers feel no desire to get into it.

I had read, at a young age, Weinstein and others.[11] At the time, I was working in the mathematical statistics department in Leiden in the Netherlands. After reading de Finetti's[12, *Chapter 1*] justification of Bayesian statistics, I had decided that I wanted to work on this topic and its foundation, expected utility. Unfortunately, the only thing my supervisor of those days, a die-hard frequentist, had ever told me was that this Bayesian approach was nonsensical and that I had better turn my attention to other topics. Fortunately, he and the academic system in the Netherlands in those days were, unlike today, liberal enough to allow me to pursue my own interests and to go in search of other supervisors. During a lunch break in Leiden, a medical statistician there, Jo Hermans, inquired about my research interests and then recommended Weinstein and others[11] to me.

I was deeply impressed by Weinstein and others.[11] This book really showed—and made tangible—that EDT can be used for practical purposes, more than any other work that I knew of. It made me more determined to pursue my Bayesian interests. In the many years that I have worked in EDT, this book has always served as a hallmark for me, showing the general applicability of expected utility. It also opened the world of MDM to me. Although the book may be technical for many MDM'ers, for me its language was ideal for making the connection with MDM. Sox and others[13] provided an alternative presentation of the basic ideas in MDM, in a language more accessible to a wide audience. I envied their way of presentation, knowing I would not be able to produce such myself.

I started working in MDM in 1992, at the MDM unit of the Leiden University Medical Center, led by Job Kievit. Job can read mathematics, economics, psychology, and medicine and is the perfect person to lead an interdisciplinary group. He could read Figure 1 and find useful things for me to do in MDM. I was lucky to live in a small country, the Netherlands, where many people are active in the field of MDM. There were, however, some surprises awaiting me, as well as some cultural shocks. I will describe 3 surprises in the next sections.

4. SURPRISE 1: ON THE TIME TRADEOFF

The time tradeoff method has been widely used to measure the quality of life of impaired health states. A typical inference from this method is as follows. Assume that we observe the following, with $\sim$ indicating that 2 things are equally preferable to some given patient:

$$\text{10 years blind} \sim \text{9 years perfect health.} \tag{4.1}$$

Then the time tradeoff (TTO) method will conclude that

$$\text{Quality of life of being blind} = 9/10 \tag{4.2}$$

for this particular patient (i.e., it is the ratio of the durations). The TTO method was one of the first machines of MDM I learned about, and I could not understand it. Shouldn't utility of life duration play a role somewhere? Any serious rational decision theory and some algebra will lead to the conclusion from equation (4.1) that, with U denoting the utility of life years,

$$\text{Quality of life of being blind} = U(9)/U(10). \tag{4.3}$$

People using TTO (equation (4.2)) apparently must have forgotten to write U, and U was dropped from the equation inadvertently. I will call this inadvertent dropping the first theoretical problem of TTO.

Equation (4.2) could be defended by arguing that usually we are not able to use U because we do not know what U is. In that case, then, we should take our best guess of U. Without further information,

$$U(y \text{ years}) = e^{-0.03y},$$

as when discounting by 3%, is a good guess and surely is better than taking U linear as in equation (4.2). Erroneously taking U linear when it should be concave (increasing at decreasing speed) leads to an erroneous overweighting of long life durations relative to short life durations. Thus, the ratio in equation (4.2) has its denominator too large and will be too low, leading to systematic underestimations of quality of life.

I was lucky in 1992 to have ended up in Leiden because Anne Stiggelbout and colleagues were carrying out empirical studies into utility corrections for TTO, such as later published by Stiggelbout and others.[14] Here I found colleagues well suited to discuss these issues with. I expected my first task in MDM in 1992 to be a simple one. I would simply explain to every MDM'er using the TTO as in equation (4.2),

"Hey you, you forgot U,"

and would hand out equation (4.3). MDM'ers would then immediately correct their mistake, and from there on they would perform TTO in the corrected manner, as in equation (4.3), for the rest of their lives. And they would pass this message on to their colleagues, so that this mistake would disappear from the scene rapidly.

Here a surprise was awaiting me, and something strange occurred, in line with my experience with Lia Verhoef in 1989. MDM'ers would not follow my advice. In the MDM meeting of 1994, the TTO was still being used as in equation (4.2), contrary to my expectations. And it still is today.

There is a second theoretical problem of the TTO method that I want to mention only briefly here. It is that the TTO method concerns intertemporal choice (trading off quality of life against life duration) and will yield utility and quality of life as relevant for intertemporal choice. In the 1980s, the common thinking in EDT was that such utility (and quality of life) need not be the same as the utility (and quality of life) relevant for risky choice (trading off quality of life against probabilities) or for interpersonal welfare considerations (trading off quality of life of one person against that of another person). Nevertheless, in cost-effectiveness studies with risk and/or welfare

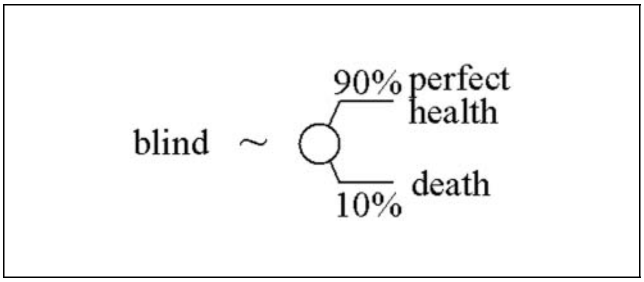

Figure 2 Standard gamble observation.

involved, MDM'ers would use the "intertemporal" scores resulting from TTO measurements without further ado. We will return to this problem later.

I now turn to the third theoretical problem for the TTO method: there is an alternative method that, at least from the theoretical perspective (judging from inside the machine), looks considerably better—the standard gamble (SG). For example, if living blind for the rest of one's life is equally preferable as a 10% risk of dying with a 90% chance of living in perfect health (with perfect sight; see Figure 2), then the quality of life of being blind according to expected utility is 0.90U(perfect health) + 0.10U (death) = $0.90 \times 1 + 0.10 \times 0 + 0.90$. I use here the common scaling convention of U(perfect health) = 1 and U(death) = 0. This method provides correct quality-of-life measurements according to the important theory of expected utility. The neat MDM expression of "gold standard" indicates this aspect of the SG method.

As is the case with the first problem, the second and third theoretical problems of the TTO have not prevented it from being widely used in MDM. This can be explained by the black-box approach of MDM, considering machines only from the outside. We will return to this point later.

5. SURPRISE 2: THE HEALTHY-YEARS EQUIVALENT

In this section and section 9, more than in other sections, I will express personal opinions that certain people in MDM may disagree with. In particular, advocates of the healthy-years equivalent method will disagree with my claims. Given the nature of this article, which is an account of my personal impressions while working in MDM, I write my opinions in nonneutral terms.

It may be desirable to develop variations of the TTO method that, like the SG, do involve risk or uncertainty. The healthy-years equivalent (HYE) method was developed for this purpose.[15] Figure 3 illustrates the method. Readers can skip the figure without loss of continuity.

It was claimed that *in theory*, the HYE measures risky utility, unlike the TTO (cf. the second problem of TTO in section 4). However, by a theoretical condition called *transitivity*, it immediately follows that Figure 3, *in theory*, measures nothing other than the TTO observation of equation (4.1). Whatever risk attitude might come in during stage 1 goes out in stage 2. Thus, in theory, HYE observations are nothing but a roundabout way of making TTO observations.[16]

In MDM, theoretically right or wrong is not considered very important. The theoretical assumption of transitivity may be systematically violated empirically. It is not impossible that alternative nontransitive theories could be developed that could reveal an interest for the HYE procedure after all. MDM'ers did not pursue such alternative theoretical justifications but took HYE simply as yet another machine, looking at it only from the outside. Many empirical investigations into the method were carried out. This was my second surprise in MDM. Testing the HYE is like testing empirically whether $2 + 2 = 5$. Given the theoretical mistakes of the inventors of HYEs, the absence of valid theoretical counterarguments to a published criticism,[16] and the absence (as yet?) of any other atheoretical justification, the HYE approach is, at best, a random idea. It should, therefore, not have received the attention it did. We will return to this issue later.

6. SURPRISE 3: ON RISKY UTILITY VERSUS RISKLESS UTILITY

I display the following statement of an intuition that has been widely shared.

> There must be more to risk attitude
> than the riskless subjective value of outcomes. (6.1)

This intuition led to an approach that was popular in EDT in the 1980s and that, in the spirit of those years, remained within the realm of expected utility.[17] It was assumed that there was a riskless cardinal utility function, often called a *value function*, that captured the intrinsic subjective value of outcomes. Then a (nonlinear) transformation was applied to convert this function into a cardinal ("risky") utility function to be used in expected

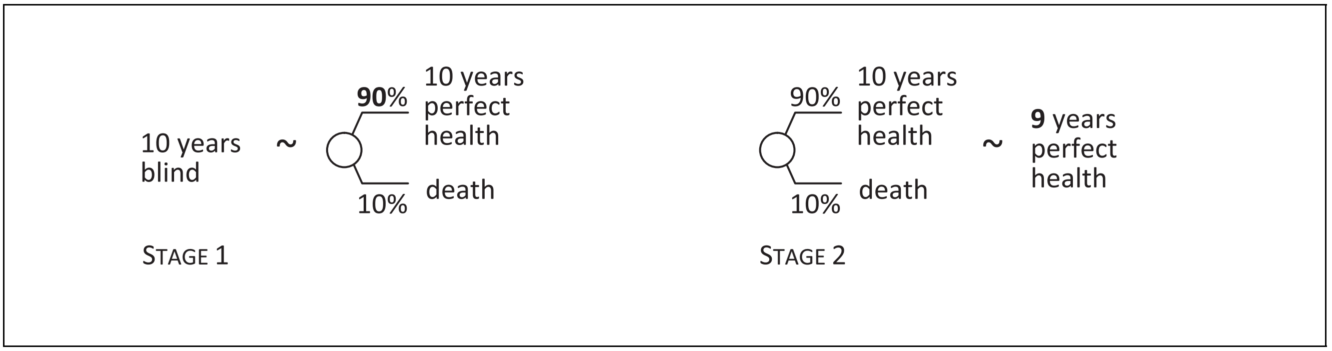

Figure 3 Healthy-years equivalent (HYE) observations. In stages 1 and 2, the bold numbers give equal preferability. For an impaired health state, in stage 1, the standard gamble question of Figure 2 is asked with a life duration specified. For the gamble resulting from stage 1, stage 2 elicits the duration in perfect health equally preferred. This duration is the HYE. In this figure, HYE(10 years blind) = 9 years.

utility calculations. This approach was usually alluded to through the expression

$$\text{Risky utility} \neq \text{riskless utility}. \tag{6.2}$$

I have always felt uneasy about the expression of riskless utility. This expression designates a very broad category, lumping together intertemporal utility as used in discounted utility, interpersonal utility as used in utilitarian welfare evaluations, and strength of preference utility to be based on introspection. These different versions of riskless utility can be as far apart from each other as from risky utility. Nevertheless, the expression of equation (6.2) was popular in EDT in the 1980s.

The status of equation (6.2) has been ambivalent in both MDM and EDT. In MDM, equation (6.2) is part of the common *thinking* on risk attitude up to the present day. At the same time, the possible distinction is commonly ignored in applications. There, utilities inferred from intertemporal (riskless) choices such as the TTO are freely used to evaluate risky evaluations and also interpersonal welfare evaluations. Utilities inferred from risky choices in SG measurements are also freely used in other contexts.

In EDT, equation (6.2) was central in the 1980s but has gradually been supplanted because prospect theory has provided a more satisfactory approach for distinguishing between the subjective value of outcomes and risk attitude. We will elaborate on this theory in section 7. New components of risk attitude different from transformed utility were introduced in prospect theory. Subsequently, risk attitude could be distinguished from utility using these new components, and equation (6.2) was no longer needed. That equation (6.2) is still commonly alluded to in MDM illustrates that this field took the ideas from EDT of the 1980s and continues to maintain these ideas at present.

It should be well understood that, to my joy, there have been many studies in MDM that did use prospect theory. Yet, its fundamentally different views on risk attitudes have not been widely understood in MDM today. I will next turn to lessons learned from the 3 surprises described in the preceding sections. In preparation, the next section briefly explains prospect theory.

7. PROSPECT THEORY

In the classical expected utility theory, risk attitude is merely captured by

COMPONENT 1 Utility of outcomes.

For example, a concave utility function of life duration implies risk aversion for life duration, and a concave utility function of money implies risk aversion for money. Many people felt that it is unsatisfactory for risk attitude to be captured by how a decision maker feels about outcomes and that it is desirable that other components be developed, such as how the decision maker feels about probability. This intuition, together with the violations of expected utility that have been observed empirically, has led to the development of new theories.[18] The most prominent new theory is prospect theory.[10,19,20]

Prospect theory was a breakthrough. Whereas previously it had been thought that irrational behavior is chaotic and cannot be modeled, this theory was the first to provide a tractable and realistic model of irrationality. It was the first rational model of irrational behavior, so to speak. Although the theory had an enormous impact on empirical and psychological studies from the beginning, at first it did not have much impact on theoretical studies in EDT. Despite its theoretical tractability, it did still have some theoretical problems, and the inside of the machine was not yet totally settled. This changed in 1992, when Tversky and Kahneman[10] were able to correct the theoretical problems using ideas from other papers.[7–9] In prospect theory, risk attitude is captured by the first component of utility plus some other components explained next:

Component 2 Probability weighting
Component 3 Loss aversion
Component 4 Framing (including scale compatibility).

Probability weighting models the ways people feel about probabilities. Typically, high probabilities are underweighted with, for example, a probability mass of 95% weighted as, say, only 80% value mass. Small probabilities are overweighted with a probability mass of 5% weighted as, say, 10% value mass. This leads to an inverse S-shaped probability transformation that has commonly been found in empirical studies. It is natural that a theory about decision under risk should not only model how people feel about outcomes but also how they feel about probabilities. Empirical studies have confirmed this route and the prevalence of the inverse S shape.[21–23] It thus is very desirable to include this component in empirical studies of risk attitudes.

Loss aversion is one of the most pronounced (although also most volatile) components of risk attitude. It reflects people's overweighting of losses relative to gains in decision making. Often, losses are weighted more than twice as much than gains of a comparable size. My guess is that more than half of the empirically observed risk aversion is generated by loss aversion. Framing, finally, combines many psychological aspects where different perceptions of the same risk, which should lead to the same decision by any rational theory, often do lead to different decisions in reality.

Prospect theory has been very successful in modeling risk attitudes. It shared the Nobel Prize in economics in 2002, and the paper by Kahneman and Tversky[19] was the second most-cited paper published in an economic journal in the last quarter of the past century.[24, TableA.10] Although the citation scores common in EDT, with its different conventions of citations and publication delays, will not always impress MDM'ers, it is hoped that this statistic will convince them that prospect theory did at least have some impact.

8. LESSON 1: ABOUT TTO

Both for EDT'ers specialized in the inner parts of machines and for MDM'ers who are best at judging machines from the outside, lessons can be learned from interactions and surprises about each other's conventions. This section discusses lessons that can be learned from the different approaches to judging the TTO method.

Prospect theory gives insights into the deviations from expected utility and, accordingly, into the errors contained in the SG method. As it turns out, all of these errors go in the same direction. One error is due to the underweighting of the good-outcome probability of gambling, a second is due to loss aversion (paying more attention to the outcome inferior to the sure option than to the outcome superior to the sure option), and a third is due to scale compatibility (because answers are expressed in probabilistic terms in SG questions, the risky nature of the gamble probabilities becomes more salient). All of these errors generate distortions in the same direction. They all worsen the evaluation of the gamble, requiring extra high probabilities there to offset the sure outcome and leading to overestimations of quality of life.

The errors for TTO measurements go in mutually opposite directions and tend to neutralize each other. Nonlinear (concave) utility of life duration augments the willingness to give up life duration. However, scale compatibility and loss aversion, highlighting the life duration given up, reduce this willingness. Hence, overall, the TTO results are less biased. Bleichrodt[25] gave a detailed account of these points, which also show that Lia's decision not to use the corrected formula in equation (4.2) was a correct one. The neutralizing effect of ("rational") nonlinear utility would have been lost after such a "correction."

With these new insights acquired into the inner parts of our machines, and with more empirical realism and understanding of irrationalities as enforced upon me by exposure to MDM, my opinion about the TTO changed. I came to think of this method as being better than the SG. Of the presently popular methods for measuring quality of life, the TTO probably is the best. Looking back at my first surprise when entering MDM in the early 1990s, as well as my early discussions with Lia Verhoef, I had to conclude that the

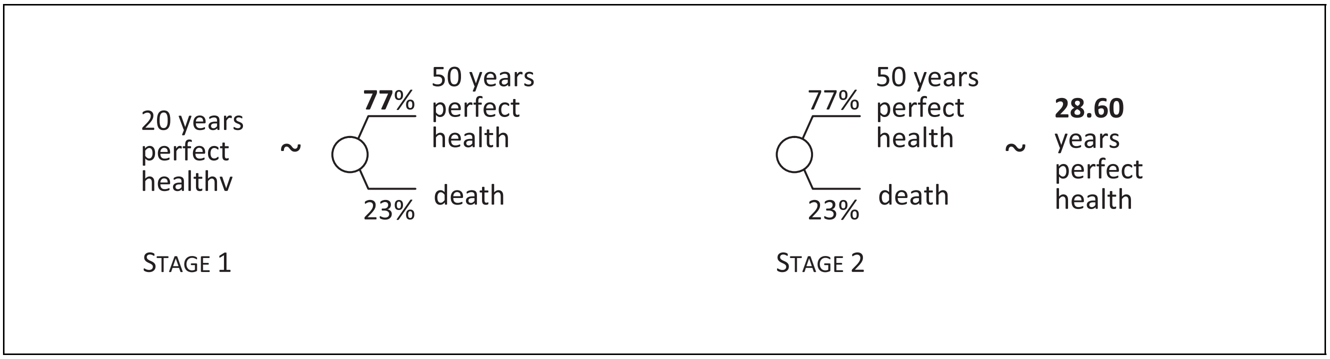

Figure 4 *Average healthy-years equivalent (HYE) observations in experiment.*

black-box approach of MDM, judging machines from the outside, had led to better conclusions in the early 1990s than my theoretically driven judgments that had been based on the insides of machines (understanding only the rationality parts and not the descriptive parts). Lia was right after all. This lesson taught me to respect and appreciate the black-box approach. Although this lesson can be stated in just a few words, I was only able to experience it properly after several years of exposure to MDM.

9. LESSON 2: ABOUT HYE

For readers who view this article as a contest between MDM and EDT, we ended the preceding section with a 1–0 lead for MDM, the outsiders. In this section, the insiders will level the score.

It is conceivable that the HYE could be a valuable contribution even if the theoretical arguments put forward to defend it are not sound. After all, theories are imperfect. However, the empirical performance of HYEs is not good. The new insights into decision theory developed in EDT in the 1990s, based on prospect theory, confirm this claim and suggest that HYEs will overestimate quality of life even more than the SG does. For those readers interested, I explain the point using Figure 3.

The first stage in Figure 3 brings all the overestimations of the SG. In stage 2, scale compatibility leads to a further bias upward: answering in terms of life duration makes the loss of life duration (9 instead of 10) more salient and the gain in certainty less so. A second bias, loss aversion, enhances this salience. Nonlinearity of the utility of life duration, which distorts TTO measurements in a neutralizing way, plays no role in the calculations of HYE. Thus, the extra complexity of HYEs not only makes the method more difficult to apply but, even worse, amplifies the distortions. Although in the SG there are already 3 biases all going in the same direction, the HYE method adds 2 more biases going in that same direction too. Now 5 biases all enhance the overestimation of quality of life.

In an informal experiment with students carried out in 1998–1999, I measured the HYEs of 20 years in perfect health through the procedure of Figure 3 (with 50 years of perfect health as best gamble-outcome). It was done in courses with groups of 8 students using forms that were numbered at the back in ways the students did not notice, so that they thought the forms were anonymous. During a break I would fill out, after the students had left the room, their individual questionnaires for the next hour. For each student, I substituted his or her individual answer of stage 1 in stage 2 in a within-subject matching ($n = 15$). The students, thinking all forms were anonymous and having answered other questions in between, were unaware of the matches. When informed later, they were highly puzzled, first about the matching that they could not explain, thinking the forms had been anonymous, and second about the discrepancies in their answers. By any sensible theory, the HYE of X healthy years should be X years, for any number X. The procedure of Figure 3, however, led to large overestimations. For within-subject matching, the average HYE of 20 years was 28.60, exceeding 20 significantly ($t_{14} = 2.78$, $P = 0.01$; see Figure 4). The HYE of 20 years of perfect health being 28.60 years of perfect health does not make sense. Similar findings were obtained with several other groups of students for 9 life durations other than 20 years and also for between-subject matchings.

My conclusion about the HYE idea is that, when it could survive for several years in the MDM community, MDM could have benefited more from inputs from EDT. In EDT, the HYE idea would not have had a chance of receiving so much attention. Here the black-box approach of MDM led to a waste of human resources. This waste could have been prevented by paying more attention to the insides of machines or to people who can judge on such insides. Such people were present in MDM,[16] but they apparently were not numerous enough to prevent the waste of human resources that took place.

10. LESSON 3: ABOUT RISKY UTILITY VERSUS RISKLESS UTILITY

The new concepts of prospect theory refer to the perception of probabilities and losses, which are conceptually more interesting than an additional layer pertaining to outcomes as in equation (6.2). The new concepts will also capture more variance in the data because they operate on dimensions other than the outcome-dimension. Hence, equation (6.2) has become obsolete in EDT. With no more need to have different cardinal utility scales, new hope could arise for a single cardinal utility scale in the social sciences that can be applied to risk, intertemporal choice, interpersonal welfare aggregations, and other contexts. This development will also resolve the second problem of the TTO method of section 4 (about using intertemporal utility in other domains). Here EDT can benefit, if not from the common theoretical thinking in MDM, then from the practice in MDM. In practice, utilities derived from one context have been freely applied to other contexts simply because this is the only way in which utility can serve as a useful tool. Here the practice of MDM has preceded the theory of both MDM and EDT, and rightly so in my view.

The ideas about utility just discussed have been elaborated by Abdellaoui and others.[26] I presented this work to economics audiences, where the relevance of introspective data for economic questions ("the allocation of scarce resources") is something like a revolutionary innovation. It was propagated by others,[27,28] with interesting recent inputs from neuroeconomics, but still mainstream economists are reluctant to consider the possibility,[29] and the revolution is yet to come. I also presented this work to MDM'ers. Here the use of introspective data for allocating scarce resources cannot be called a novelty because MDM'ers do it themselves every day, such as in the so-called visual analog scale (VAS) measurement of quality of life.[11,13,30] For MDM audiences, I would emphasize the novelty that introspective data can now be related to the decision foundations of economics, proving their relevance for gold standards of rational choice. The study by Abdellaoui and others,[26] written for an economics audience, referred to Stalmeier and Bezembinder[31] as a predecessor in the medical literature.

11. DISCUSSION OF DESCRIPTIVE PROSPECT THEORY AND NORMATIVE EXPECTED UTILITY

Quality-of-life measurements usually concern persons from the general public and patients who are interviewed without having had many learning or training opportunities and with questions that usually concern hypothetical decisions. In such measurements, the effects described by prospect theory will arise prominently, and reanalyses of empirical findings in MDM in terms of prospect theory will be worthwhile. I want to emphasize here that prospect theory is purely descriptive. Behavior will move toward the rational expected utility model if subjects can think deeply about their decisions and have the chance to learn.[19,32–35]

If quality-of-life measurements can be interactive with much opportunity for learning, then it may be desirable to move in the direction of rational behavior and expected utility. Then violations of expected utility can be reconciled interactively, which is preferable to the "mechanical" solutions based on prospect theory whenever possible.[36–39]

At the stage of decision making where the data, including quality-of-life assessments, are available and an optimal policy has to be chosen or recommended, we have arrived at the prescriptive realm of optimal decision making. Then expected utility is the best theory to determine which decisions to take.

12. CONCLUSION

Different habits and customs in medical decision making and economic decision theory have led to different views on the measurement of quality of life and utility, where mutual benefits result from interactions. Combining insights from both fields leads to a number of suggestions for improvements. Based on prospect theory, from the currently conventional decision-based methods for measuring quality of life, the TTO method is best because its biases neutralize each other. For the future of quality-of-life measurement, the TTO method can, however, not be the final answer. It will be desirable, for instance, to have risk included and not (just) intertemporal tradeoffs, as is the case for TTO. Whereas the various biases for the TTO method may neutralize each other, they still are present and do generate noise. Thus, future versions of the SG may be developed that are not subject to the many biases to which the current measurement methods are. Providing more opportunities for learning, training, and motivation ("getting more out of fewer subjects") is a promising direction.

REFERENCES

1. McNeil, BJ, Weichselbaum R, Pauker SG. Fallacy of the five-year survival in lung cancer. N Engl J Med. 1978;299:1397–401.

2. Keeney RL, Raiffa H. Decisions with Multiple Objectives. 2nd ed. Cambridge, UK: Cambridge University Press; 1993.

3. Krischer JP. An annotated bibliography of decision analytic applications to health care. Oper Res. 1980;28:97–113.

4. Arrow KJ. Alternative approaches to the theory of choice in risk-taking situations. Econometrica. 1951;19:404–37.

5. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions: health utilities index. PharmacoEconomics. 1995;7:503–20.

6. Hammond KR. Beyond Rationality. New York: Oxford University Press; 2006.

7. Gilboa, I. Expected utility with purely subjective non-additive probabilities. J Math Econ. 1987;16:65–88.

8. Quiggin J. A theory of anticipated utility. J Econ Behav Organ. 1982;3:323–43.

9. Schmeidler D. Subjective probability and expected utility without additivity. Econometrica. 1989;57:571–87.

10. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncert. 1992;5:297–323.

11. Weinstein MC, Fineberg HV, Elstein AS, et al. Clinical Decision Analysis. Philadelphia: W. B. Saunders; 1980.

12. de Finetti B. Probability, Induction and Statistics. New York: Wiley; 1972.

13. Sox HC, Blatt MA, Higgins MC, Marton KI. Medical Decision Making. Stoneham, MA: Butterworth-Heinemann; 1988.

14. Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Stoter G, de Haes JCJM. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. Med Decis Making. 1994;14:82–90.

15. Mehrez A, Gafni A. Healthy years equivalents: how to measure them using the standard gamble method. Med Decis Making. 1991;11:140–6.

16. Johannesson M, Pliskin JS, Weinstein MC. Are healthy-years equivalents an improvement over quality-adjusted life years? Med Decis Making. 1993;13:281–6.

17. Dyer JS, Sarin RK. Relative risk aversion. Manage Sci. 1982; 28:875–86.

18. Starmer C. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. J Econ Lit. 2000;38:332–82.

19. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica. 1979;47:263–91.

20. Wakker PP. Prospect Theory: Rank and Sign Dependence for Risk and Ambiguity. Cambridge, UK: Cambridge University Press; in press.

21. Abdellaoui M. Parameter-free elicitation of utility and probability weighting functions. Manage Sci. 2000;46:1497–512.

22. Bleichrodt H, Pinto JL. A parameter-free elicitation of the probability weighting function. Manage Sci. 2000;46:1485–96.

23. Gonzalez R, Wu G. On the shape of the probability weighting function. Cognit Psychol. 1999;38:129–66.

24. Coupé T. Revealed performances: worldwide rankings of economists and economics Departments, 1990–2000. Working paper. Available from: URL: http://homepages.ulb.ac.be/~tcoupe/updaterevealedperformances.pdf

25. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health Econ. 2002;11:447–56.

26. Abdellaoui M, Barrios C, Wakker PP. Reconciling introspective utility with revealed preference: experimental arguments based on prospect theory. J Econometrics. 2007;138:336–78.

27. van Praag BMS. Individual Welfare Functions and Consumer Behavior. Amsterdam: North-Holland; 1968.

28. Kahneman D. New challenges to the rationality assumption. J Instit Theor Econ 1994;150:18–36.

29. Gul F, Pesendorfer W. The case for mindless economics. In: Caplin A, Schotter A, eds. Foundations of Positive and Normative Economics. New York: Oxford University Press; 2008.

30. Mortimer D, Segal L. Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. Med Decis Making. 2008;28:66–89.

31. Stalmeier PFM, Bezembinder TGG. The discrepancy between risky and riskless utilities: a matter of framing? Med Decis Making. 1999;19:435–47.

32. List JA. Neoclassical theory versus prospect theory: evidence from the marketplace. Econometrica. 2004;72:615–25.

33. Plott CR. Rational individual behaviour in markets and social choice processes: the discovered preference hypothesis. In: Arrow KJ, Colombatto E, Perlman M, Schmidt C, eds. The Rational Foundations of Economic Behavior: Proceedings of the IEA Conference Held in Turin, Italy. New York: St. Martins; 1996. p 225–50.

34. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science. 1981;211:453–8.

35. van de Kuilen G, Wakker PP. Learning in the Allais paradox. J Risk Uncert. 2006;33:155–64.

36. Kimbrough SO, Weber M. An empirical comparison of utility assessment programs. Eur J Oper Res. 1994;75:617–33.

37. March JG. Bounded rationality, ambiguity and the engineering of choice. Bell J Econ. 1978;9:587–608.

38. Raiffa H. Risk, uncertainty and the Savage axioms: comment. Q J Econ. 1961;75:690–4.

39. Slovic P. The construction of preference. Am Psychol. 1995; 50:364–71.