Title: The sequence effect in time trade-off elicitation.

Authors: Michaël Schwarzinger, MD, MPH, Thea Herz, PhD, Isabelle Durand-Zaleski, MD, PhD.

Departments and institutions involved in the work: Institut National de la Santé et de la Recherche Médicale, INSERM unit 444 (MS), Faculté de Médecine Saint-Antoine, Paris; Service de Pharmacologie Clinique (MS), Hôpital Henri Mondor, AP-HP, Créteil; Laboratoire de Psychologie Sociale, Université René Descartes (TH), Paris; Service de Santé Publique (IDZ), Hôpital Henri Mondor, AP-HP, Créteil, France.

The paper was presented at the 8th biennial meeting of the European Society for Medical Decision Making, Taormina, Sicily, Italy, June 2-5, 2002.

Address correspondence and reprint requests to Dr. Schwarzinger: Service de Pharmacologie Clinique, Hôpital Henri Mondor, 51 avenue du Maréchal de Lattre de Tassigny, 94010 Créteil, France; e-mail: schwarzi@univ-paris12.fr
Abstract

Background: Most valuation experiments include only a few health states related to a specific condition. We tested the sequence effect in time trade-off elicitations, i.e. the value given to a particular health state would be higher when the previous health state elicited was a milder health state as compared to a worse health state. Methods: 54 nurses enrolled prospectively were randomly assigned to one of two sequence valuation strategies: a 'mild-to-moderate' sequence from chronic low back pain to asthma, or a ‘severe-to-moderate’ sequence from permanent impairments after stroke to the previous asthma. Results: Time trade-off values of the initial disease stages were 0.87(std 0.18) for chronic low back pain and 0.45(std 0.12) for moderate permanent impairments after stroke. The value of severe asthma was significantly higher in the ‘mild-to-moderate’ than in the ‘severe-to-moderate’ evaluation sequence (0.84 vs. 0.72, p<0.04).

Conclusion: The sequence effect revealed in our experiment is in line with other findings showing framing effects and violations of procedure invariance in health state valuation.

Keywords: procedure invariance; sequence effect; order effect; time trade-off; QALY; DALY
Introduction

Quality-Adjusted Life Years (QALYs) take into account both the number of years lived in a particular condition and the relative severity of a year lived in that condition. QALYs have been recommended to measure effectiveness of health interventions in cost-effectiveness analysis.\(^1\) The number of published cost-utility analyses relying on QALYs increased exponentially with time.\(^2\) However there is no gold standard method to elicit the relative severity of year lived in a particular condition. Various valuation methods have been developed following rational theories of riskless or risky choice (time trade-off and standard gamble, respectively).\(^3\) Whatever the valuation method, respondents are expected to maximize some subjective measure associated with each health state and thus to reveal their preferences for health states.

A basic assumption of rational theories of choice is the principle of invariance. It states that the relation of preference between health states should not depend on the description of the options (framing invariance) nor on the method of elicitation (procedure invariance). Without stability across equivalent descriptions and equivalent elicitation procedures, respondents' preferences cannot be represented as maximizing some subjective measure. A few experiments have addressed this issue in health economics. Within a given valuation method, framing invariance was not supported by various types of health state description,\(^4 - 7\) nor when reference anchor varied from perfect health to death.\(^4, 7 - 9\) Procedure invariance was not found in time trade-off and standard gamble methods when respondents completed a written questionnaire or were interviewed with props\(^10\) or by means of a computer,\(^11\) nor when usual ping-pong search procedure was challenged by a top-down titration.\(^12\)
Most valuation experiments include only a few health states related to a specific condition. We hypothesized that the value of a particular health state is not independent from the health state valued previously, e.g., the value given to a particular health state would be higher when the previous health state was a milder health state as compared to a worse health state. In this paper, we looked for this sequence effect in time trade-off elicitations with a randomized trial design.
Material and Methods

Study design

We defined two sequence valuation strategies. In the ‘mild-to-moderate’ sequence valuation strategy, participants valued successively a mild disease stage, i.e. chronic low back pain, a moderate disease stage related to asthma, and the mild disease stage again. In the ‘severe-to-moderate’ sequence valuation strategy, participants valued successively a severe disease stage, i.e. moderate permanent impairments after stroke, the previous moderate reference disease stage related to asthma, and the severe disease stage again. Participants were randomly assigned to either sequence valuation strategies. Our main hypothesis was that the value of the reference moderate disease stage would be higher in the ‘mild-to-moderate’ than the ‘severe-to-moderate’ valuation strategy, i.e. a sequence effect.

Population

We selected an homogenous group regarding both knowledge of health conditions elicited and willingness to participate in valuation experiments. Thus we enrolled prospectively French nurses following a MPH course during the spring 2001 (Ecole Montsouris, Paris).

Disease stages' descriptions

The disease stages’ descriptions were taken among 28 disease stages covering 13 diseases and validated in the European Disability Weights project by external health care professionals and public health experts. All disease stages were described on a sheet with the name of the disease at the top, the position of the selected disease stage among the other stages, a brief clinical
description and a health state profile, i.e. EQ-5D extended with a cognitive dimension.\textsuperscript{15-17} We selected the three disease stages based on their individual unambiguous ranking in the European Disability Weights project:\textsuperscript{14} chronic low back pain, severe asthma, moderate permanent impairments after stroke received overall time trade-off values of 0.87, 0.80, and 0.58, respectively (n=232). Their presentation is shown in Figure 1.

Valuation method

After a course on health economics, nurses were asked to participate in a short elicitation exercise. They were interviewed by a single interviewer (MS) following a standardized elicitation protocol. After presentation of the first health state (either chronic low back pain in the ‘mild-to-moderate’ sequence valuation strategy, or moderate permanent impairments after stroke in the ‘severe-to-moderate’ sequence valuation strategy), a time trade-off was used with the following features: a fixed 40 years life-expectancy; a fixed gain formulation (e.g. ‘do you prefer to live 30 years in full health over living 40 years in the described health state?’); and a ping-pong elicitation procedure starting by the extreme anchors (i.e., full health and death) to check for understanding.

Statistical analysis

Disease stage values were computed on a scale from zero (death) to unity (full health) by dividing by 40 the number of years for which respondent was indifferent. With the risk of a type I error at 5 percent (one-sided test) and the risk of a type II error at 20 percent, the required sample size was 50 nurses to show a sequence effect of 0.1 on the mean value of the reference moderate disease stage (with standard deviation of 0.1). We used the Student t test to compare
the value of the reference moderate disease stage between the ‘mild-to-moderate’ and the ‘severe-to-moderate’ groups. Other binary data were compared by a chi-square test or Fisher's exact test. Significance was attributed at the 5% level, and data were analysed with SAS 8.0 (SAS Institute, Cary NC).
Results

Fifty-four nurses were approached to participate in the elicitation experiment. After the presentation of the first health state to be valued, all nurses showed a correct understanding of the time trade-off method as indirectly shown by their answers to the trade-off with reference anchors. However, one third of the nurses were unwilling to trade-off years in the ping-pong elicitation questions (41% in the ‘severe-to-moderate’ strategy group and 26% in the ‘mild-to-moderate’ strategy group, p=0.25). For the 36 nurses who were willing to trade-off years, mean age (40.3 (std 5.4)) and gender (80% females) did not differ significantly between the two valuation strategy groups.

Table 1 shows that the valuation pattern of the three health states followed the predicted pattern from the European Disability Weights project with increasing severity from chronic low back pain, severe asthma, to moderate permanent impairments after stroke. Our main hypothesis about sequence effect in time trade-off elicitations was supported by our experiment, i.e. the mean value of severe asthma was significantly higher in the ‘mild-to-moderate’ valuation group (0.84 (std 0.2)) than in the ‘severe-to-moderate’ valuation group (0.73 (std 0.18)) (t\textsubscript{35} =-1.80, p<0.04). The valuation group effect remained significant when adjusting on age and gender in ANOVA.

Finally, there was almost no change in the value given to the initial health state when it was elicited again at the end of the elicitation exercise (0.88 in chronic low back pain and 0.46 in moderate permanent impairments after stroke).
Discussion

QALYs are increasingly used in cost-effectiveness analysis and are recommended by the Panel on Cost-Effectiveness Analysis in Health and Medicine since 1996.\textsuperscript{1} However, the question remains whereas the 'Q' is simply revealed by elicitation exercises or actually constructed during the process of elicitation. The former approach relies on strong assumptions, i.e. individual preferences are well formed, complete, and stable.\textsuperscript{18,19} As shown by our study, the value of a second health state in an evaluation sequence was affected by the sequence effect with a significantly higher value in the 'mild-to-moderate' than in the 'severe-to-moderate' health state valuation strategy. The sequence effect revealed in our experiment is in line with other findings showing framing and procedure invariance in health state valuation.\textsuperscript{4-8,10-12}

To the best of our knowledge, our experiment is the first to show the sequence effect in health economics. Had we used a valuation strategy comprising only two health states, results would have signified an order effect. That is, the values given to states A and B depend on the evaluation sequence, A-B or B-A. The only study dealing with order effects in health state valuation actually controlled for order effects rather than studied them.\textsuperscript{20} In that study, 55 patients with early-stage breast cancer had to value their actual health state and a radiotherapy scenario at different times of their disease (before, during and after radiotherapy). There was no significant order effect but it may be explained by the clear overlap of the two health state values (e.g. 0.94 and 0.89 before radiotherapy, respectively). Outside health economics, order effects in preference judgments were evidenced by willingness-to-pay for options previously shown to produce joint-versus-separate preference reversals.\textsuperscript{21}
An intuitive explanation of the sequence effect is that respondents do not have any clue of health state values a priori, and that they construct their values during the process of elicitation. In order to prepare her first health state elicitation, the respondent has to generate a representation for the health state considered. When valuing the next health state, she uses the same representation, “considering” it is still relevant. Several indirect arguments could sustain this explanation: the health state elicited first in the sequence valuation strategy (i.e. chronic low back pain in the ‘mild-to-moderate’ sequence or moderate permanent impairments after stroke in the ‘severe-to-moderate’ sequence), differs not only by health-related quality of life dimensions but also the vital prognosis; respondents took a longer time to value the first health state than the second one; and there was almost no change in the value given to the initial health state when it was elicited again at the end of the elicitation exercise.

The sequence effect may be seen as an extension of the well-known anchoring effect, where response is biased toward any value, even if it is arbitrary, that the respondent is induced to consider as a candidate answer. In our experiment, respondents anchor themselves on their first estimation of health state value without any anchor provided by the interviewer. Several theories have been proposed to explain anchoring bias, such as anchoring and adjustment, and the selective accessibility model. Future research could look for feasible debiasing strategies in the sequence of health state valuation even if previous experiments showed the robustness of anchoring effects in other economic areas. It would also be of interest to compare the sequential process of valuation described in our experiment with an end-of-sequence process of valuation where all health states would be presented and ranked initially, as psychometric
experiments showed that relative ratings on a Visual Analogue Scale depend on the set of stimuli proposed to the respondent.\textsuperscript{26}

We found far from easy to offer a theoretical framework that accounts for the sequence effect. According to Prospect theory,\textsuperscript{8,27} respondents evaluate outcomes as changes from a reference point, and they value generally losses as more serious than equivalent gains. In our experiment, if respondents had taken the first health state as a reference point, then they should have given more weight to the loss of the ‘mild-to-moderate’ sequence than to the gain of the ‘severe-to-moderate’ sequence of valuations. However our results showed a reversed pattern with the value of the moderate health state pulled “up” toward the mild health state in the former sequence and “down” toward the severe health state in the latter sequence. A plausible explanation is that our moderate health state was too close to the mild health state as the absolute value difference in the ‘mild-to-moderate’ sequence (0.04) was much smaller than in the ‘severe-to-moderate’ sequence (0.28).

The sequence effect found in our experiment should be retested with other moderate health states distributed over the range of values between the mild and the severe health states to check if the sequence effect was related to the particular health states used in our study or if it is a general phenomenon. In addition, predictions from Prospect theory could have been checked if we had elicited all three health states in both sequence valuation strategies with, e.g., the expected value of the severe health state in the ‘mild-to-moderate' evaluation sequence lower than in the ‘severe-to-moderate' sequence.

Our study has some limitations. We found an unexpectedly high refusal rate in willingness to trade longevity for health quality in our French convenient sample who had nonetheless both a
high level of education and the knowledge of the conditions elicited. We decided to exclude from analysis those nurses who refused to trade-off years of life as it is usually done in valuation experiments. However, the sequence effect would have become not significant if we had imputed a value of 1 for those missing data, i.e. the mean value of severe asthma would be 0.88 and 0.84 in the ‘mild-to-moderate’ and the ‘severe-to-moderate’ valuation groups, respectively (t_{55}=-0.87, p=0.19). Of particular interest is the higher refusal rate in those presented with the severe health state first (41% vs. 26%). It may be related to the unwillingness to trade-off survival and quality of life in those who suffer from severe and life-threatening diseases. This phenomenon was found in patients affected by severe conditions,\textsuperscript{28} and in individuals eliciting hypothetical scenarios as in our experiment.\textsuperscript{29} Follow-up discussions were consistent with this phenomenon. In particular, many nurses said that they would not trade any year of life as they have young children. This high refusal rate to trade-off years of life should be replicated in a second study with a different sample, e.g., with unmarried nurses with no children or older women whose children have grown. Alternatively, Chapman et al. suggested the use of an impersonal elicitation questionnaire to avoid this phenomenon.\textsuperscript{28} More broadly, the very few cost-effectiveness analyses that relied on real French valuation experiments may reflect difficulties to translate the Anglo-Saxon QALY concept into the French health-care system (in fact only two over 125 studies collected in the French cost-effectiveness analyses database).\textsuperscript{30}

We found that the sequence effect could significantly affect the results of a health state valuation experiment. Indeed the value of severe asthma decreased by 0.11 points on a 0-1 scale from the ‘mild-to-moderate’ sequence to the ‘severe-to-moderate’ sequence valuation strategy. This new procedure effect could be added to other procedure effects,\textsuperscript{10-12} framing effects,\textsuperscript{4-8} logical
inconsistencies,\textsuperscript{31,32} or internal inconsistency,\textsuperscript{33} found in health state valuation methods that all show that health states values are poorly constructed values. The recognition of these failures of invariance should contribute to a new conception of judgment and choice in which preferences are often constructed, not merely revealed, in the process of elicitation, and their measure could follow a ‘building code’.\textsuperscript{34} Otherwise, the sequence effect may contribute to the difficulties in comparing constructed values from various studies addressing different conditions or even within a single condition at different stages.\textsuperscript{35,36}
Acknowledgments

We would like to thank Pr Arthur Elstein for useful suggestions.

References


Table 1: Valuation of three disease stages in the 'mild-to-moderate' and the 'severe-to-moderate' valuation groups, mean(std)

<table>
<thead>
<tr>
<th>Disease stage</th>
<th>Mild-to-moderate</th>
<th>Severe-to-moderate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic low back pain</td>
<td>0.88 (0.18)</td>
<td>NA</td>
</tr>
<tr>
<td>Severe asthma</td>
<td>0.84 (0.20)</td>
<td>0.73 (0.18)</td>
</tr>
<tr>
<td>Moderate permanent impairments after stroke</td>
<td>NA</td>
<td>0.45 (0.12)</td>
</tr>
</tbody>
</table>

NA: Non applicable.
Stroke

Acute incident plus rehabilitation phase
Mild permanent impairments after stroke

Moderate permanent impairments after stroke

Severe permanent impairments after stroke

Patient with permanent impairments in movement, speech and memory after incomplete recovery from a stroke more than one year ago.

- Some problems in walking about
- Some problems with washing or dressing self
- Some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- Moderate pain or discomfort (in this case, discomfort)
- Moderately anxious or depressed
- Some problems in cognitive functioning (e.g. memory, learning ability, concentration, comprehension)
**Asthma**

Mild asthma

Moderate asthma

**Severe asthma**

*Patient who has at least two attacks of severe shortness of breath a year, that each confine them to their home for one week or may lead to hospitalisation, despite adequate therapy. Between attacks the patient has limited exercise capacity, needs to take daily preventive medication and has to avoid provocative situations (e.g., smoke, dust and strong perfumes).*

**Health status outside attacks:**

- Some problems in walking about
  - No problems with washing or dressing self
- Some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- Moderate pain or discomfort (*in this case, discomfort*)
- Moderately anxious or depressed
- No problems in cognitive functioning (e.g. memory, learning ability, concentration, comprehension)
Low Back Pain

Acute low back pain

Chronic low back pain

Patient with radiating pain low in the back, limited in sitting and to a smaller extent in walking; patient is generally restricted in all physical activities; standard pain relieving treatment is provided.

- Some problems in walking about
- Some problems with washing or dressing self
- Some problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
- Moderate pain or discomfort (in this case, pain)
  - Not anxious or depressed
- No problems in cognitive functioning (e.g. memory, learning ability, concentration, comprehension)