

Work In Progress – Please do not quote without authors' permission

**A qualitative investigation to explore the validity of the visual analogue scale,  
time trade off and person trade off techniques**

Authors: Suzanne Robinson, Tim Freeman, Stirling Bryan

Health Services Management Centre  
University of Birmingham  
Park House  
40 Edgbaston Park Road  
Birmingham B15 2RT  
Tel: 0121 414 3011  
Email: [s.m.robinson@bham.ac.uk](mailto:s.m.robinson@bham.ac.uk)

## **Introduction**

Preference elicitation techniques, such as the Time Trade off (TTO), Standard Gamble, and Person Trade Off (PTO), are a key component in the calculation of Quality Adjusted Life Years, with different techniques often yielding different results (Brazier et al, (1999); Read et al, (1984); Dolan et al 1996). Applying different utility weights in a cost utility analysis could lead to different resource allocation decisions. Therefore, one of the areas of methodological interest is around the validity of the different techniques, that is do respondents do what we think they are doing when they participate in a health state valuation exercise. The QALY approach is based on a utilitarian framework that which assumes that individuals are utility maximisers and that their preferences satisfy the axioms of Expected Utility Theory (EUT), which is a normative theory of decision making, that describes how people should behave under uncertainty and not how they do behave.

EUT is based on the work of Von Neumann-Morganstern (1944) and is an approximation to behaviour under uncertainty which assumes that individuals make choices about different options (here health states) in a mathematically convenient and consistent manner, i.e. they are logical {McGuire and Drummond 2001).

Basically, the theory assumes that individuals have fully formed, highly articulated preferences which can be applied to any form of decision making. However, there is suggestion that rather than having well articulated preferences, individuals use heuristics (cognitive shortcuts) to simplify the decision making process (Kahneman, D et al (2002); Lloyd A, (2003); Cairns et al, (2002); Gigerenzer et als (1996); Payne et al (1993)). There are many factors which can influence an individual's preference including: the framing of the question; format of the session; complexity of the task and additional information 'baggage' brought by respondents. The main research to date has explored quantitative issues around reliability and convergent validity (Torrance (1976); Brazier et al, (1999)). There is less research around the cognitive process which respondents undertake in order to reach their valuations, and the meaning respondents place on their valuations. Given this, qualitative methods were used to provide a greater understanding of the process individuals use and the experience they face when participating in health state valuation exercises.

## **Methods**

### *Research design*

The valuation work was conducted in five panel sessions which followed a standardised protocol. Participants were health and non health care professionals (including General Practitioners (GPs)) the groups were not mixed in the panel sessions. All Sessions were guided by a trained facilitator and the process was designed to enhance discussion and deliberation (see Appendix 1 for further details).

Whilst statistical techniques can highlight some problems in terms of reliability and comparability between different methods they do not shed light on the way respondents interpret or answer the question. Qualitative methods are widely used in other disciplines such as psychology and sociology to explore the language of the valuation technique and cognitive process respondents undertaken when completing the task (Pope and Mays, (2000); Marshall and Rossmans, (1999); Mallinson, (2002)). It was therefore felt that a predominantly qualitative approach was appropriate for addressing the research question, given the need to explore issues around the cognitive process and how respondents complete health state valuation exercises. The qualitative investigation reported in this paper was conducted post panel session.

### *The interview schedule*

Unstructured and semi-structured interviews allow the researcher to 'understand the world from the subject's point of view, to unfold the meaning of experiences, to uncover their lived world' (Kvale S 1996). In this study semi structured interviews were used because they allow for a balance between free flowing and directed conversation. Thus, allowing particular research questions to be addressed fully giving respondents the chance to explore other issues and concerns (Berg, 1998).

The interview schedule developed in consultation with an experienced qualitative researcher, covered themes which had emerged from the literature and discussions during the health state valuation group sessions. These themes were explored during the interviews. The interview schedule was used to ensure that all of the themes were covered with each participant, and the ordering changed to follow the natural 'flow' of discussion required for rapport. 'Interviews sort to draw on respondent's experiences, opinions, beliefs, feelings, knowledge and perceptions' (Patton M, 1987).

The main question for exploration was around the process of eliciting valuations and how respondents make their choices. In order to explore this four overarching themes explored in the interview were: stimuli presented; context and framing issues; discussion and deliberation and use of values in decision making.

A brief outline of the interview schedule is outlined in box 1 (below) and the complete interview schedule is available from the authors. Various stimulus were used to aid the interview process, these included props from the earlier panel sessions and closed questions around the difficulty associated with each technique. The closed ended questions produced quantitative data that was explored in the interviews for example, if a respondent said they found a technique difficult the interviewer then went on to explore this further with the participant.

The particular questions used in this study were taken from emerging themes and discussion in the existing literature. The health economic literature has tended to use quantitative approaches that focus on the theoretical and empirical issues relating to the measurement of individual preferences, very few studies have explored the nature of these preferences (Dolan 2000). Many of the quantitative investigates have 'thrown up' further questions which are better explored in a qualitative manner i.e. 'why do responses differ in addition to how they differ' (Dolan 2000). Whilst, there is limited qualitative studies much of the existing literature does recognise the need for more qualitative investigation and suggests this as an area for future research (Dolan, 2000). This study uses qualitative investigation as a parallel tool to investigate the validity of the different valuation techniques and explore some of the emerging themes.

#### *Sample and data collection*

Purposive sampling technique was used to contact those who had taken part in the panel session. Letters were sent to all individuals who had taken part in the panel sessions, inviting them to participate in a semi structured interview. Non-responders were then contacted by telephone. All interviews were conducted by the same interviewer and tape recorded after assurances of confidentiality were provided. Transcription of tapes was undertaken independently by an individual who had no involvement with the research.

**BOX 1: A brief outline of the interview schedule is outlined below:**

The interview began with open questions, followed by more focused closed ended questions.

**Stage 1:** The interviewer explained the interview process and stressed that the interview was separate from the EDW project.

**Stage 2:** VAS method was reviewed using props from the panel session. These included the stimuli presented for valuation (see Appendix \*), the VAS instructions and thermometer (see appendix \*). Respondents were then asked to explain the process they used to reach their VAS scores.

*Questions explored included:-*

Aspects of the stimuli (disease card) focused on by respondents;  
how respondents used the numbers on the VAS;  
could respondents disregard future prognosis, economic aspects and effects on carers, when they valued health states.

**Stage 3:** PTO method was reviewed using props from the session. Respondents were then asked to review the PTO method and explain the process they used to reach their valuations.

*Questions explored:-*

Aspects of the stimuli focused on by respondent;  
framing of the question and the use of quadriplegia as an anchor state;  
hypothetical scenarios and assumptions used, how did respondents respond to the societal perspective?

**Stage 4:** TTO method was reviewed using props from the session. Respondents were then asked to review the TTO method and explain the process they used to reach their valuations.

*Questions explored:-*

Framing of the question and its effect on the valuation;  
context and perspective of question;  
Could respondents disregard future prognosis etc?  
Did respondents trade and why?

**Stage 5:** Close-ended questions were used to ask respondents the level of difficulty they associated with each technique, the answers ranged from easy – extremely difficult. Respondents were also asked to explain their answers.

**Stage 6:** This stage of the interview explored the effects of group discussion and the process of deliberation. Respondents were asked to discuss their feelings on the panel process. The interviewer also tried to explore whether the group session had any influence on the valuation given.

**Stage 7:** Respondents were asked how they felt about their valuations and their views on the use of these valuations by others in a decision-making context?

### *Data analysis*

Data analysis consisted of a modified analytic induction, in which thematic categories identified in the previous literature were used as sensitising devices in developing the analysis (Silverman 2000). These draft codes were modified and extended to ensure that all of the available data could be accounted for, the analysis thus builds on available prior themes yet is free to extend beyond that. The data then changes the codes as it identifies gaps in the existing theory and begins to develop new and emergent themes and categories. The approach used in this study

drew on the work of Miles and Humberman (1994), to ensure that a systematic refining of data analysis and the development of inductive categories in an ongoing iterative process of interpretation, essentially of dialogue between prior categorise and the available data to develop a contribution beyond that previously available. Particular attention was paid to 'negative cases' that did not fit the emergent frameworks and have moved the analysis forward, in order to show a more measured, credible analysis and increase internal validity.

Data coding followed the convention of Miles and Humberman (1994), in which around a third of the available data was taken and common ideas and emergent themes, based on both prior work and from within the data, were written on separate sheets of paper as a provisional coding scheme. This process was continued for the next third of the data and categories were further refined and developed, following modifications, the final third of data was analysed and the coding frame further refined. Once finalised the coding frame was applied to all of the data, and all data relating to each of the codes brought together for further analysis, in which patterns and paradoxes were explored. This final stage was done using NVIVO which is a data handling software package.

#### *Data reporting*

The analysis reported below used exemplar quotations from interviews to illustrate findings as well as some summarising and interpretation. As the main purpose is thematic, no descriptive quantitative analysis is given. However, where appropriate some simple quantification of findings were presented relating to information collected from the closed ended questions used in the study, such as responses to the level of difficulty associated with the techniques.

#### **Results**

A total of 20 interviews were undertaken including 2 health care professionals, 18 non health care professionals all from the University of Birmingham. The fact that respondents had been involved in a panel session that lasted around 5 hours meant that many were reluctant or could not afford to give up extra time to be interviewed. This was especially apparent for GPs none of whom were available for interview.

#### *Stimuli presented*

The tools used by health economists to measure individuals preferences are based on the notion that people have 'fully formed and highly articulated preferences which they can quickly and accurately access and apply to any form of decision making' (Loomes G 1998). This method

assumes that individuals are able to process all the information presented. In terms of the stimuli presented here it means that respondents would be able to process all the information and trade between the attributes (disease label and generic measure) in order to make their decision and maximise their utility. However, there is evidence in the literature to suggest that respondents do not engage in compensatory decision making, instead they are motivated to simplify the decision making task (referenced by Lloyd A, (2003); Payne et al (1993)) and the more complex the task becomes the more likely they are to resort to simplifying heuristics.

In this study respondents were asked to value health states which had both a disease label (including the disease name and a description of the disease state) and a generic functional health status EQ-5D+ 1 (see box 2) The interviews attempted to establish what aspects of the disease card respondents had focused on in the sessions and to explore whether respondents engaged in compensatory decision making.

Box 2: Example disease-specific scenario

<b>Vision Disorder</b>	<b>SVIS</b>
Mild /moderate vision <b>Severe vision disorder</b> ←	

**Patient is unable to read small newspaper print and has great difficulty or is unable to recognise faces at 4 meters distance**

- No 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)
- No pain or discomfort
- Moderately anxious or depressed
- No problems in cognitive functioning (e.g. memory, learning ability, concentration, comprehension)

*Disease label, functional health status or both?*

The data from the interviews suggest that respondents focused on different aspects of the stimuli depending on the complexity of the question and the health state being valued.

For example, in terms of the VAS if the health state could easily be ranked to best or worst imaginable health state, respondents would just look at the disease label. In contrast, if there were a group of health states that were clustered around a similar point then respondents focused more on the functional health status.

*“I used all of it (points to disease card). But I think it was, when it came down to, there was sort of different degrees. I mean for those where, as I say, it felt self-evident that they*

*were bad or the worst, you tended to focus on the statement, you know like quadriplegia that's bad so I don't need to read the rest. But then when you were actually, some of the where you weren't certain where to put them, it was much more about the functional, you know, what does it actually mean in practice? How many dots are there" (ID 9: panel 2).*

For this particular respondent it would seem that they are engaging in compensatory decision making, in so far as they are processing all the information and trading between the different attributes. However, a common theme which emerged in the interviews was the dominance of the disease quadriplegia which was referred to by most respondents.

One respondent suggested she compared everything to quadriplegia which was for her the worst imaginable health state.

*"I intermingled information on the card, because as I say, the one that probably did keep leaping out at me, was spinal cord, I just don't think I could cope with being a paraplegic or whatever, and having to be confined to bed and have everything done for me, so that was always my focus, and that probably might have clouded me looking at anything else, because that was the worst and I just compared everything to that" (ID 7 panel 1).*

On interpretation of this behaviour is that respondents are using *fast and frugal heuristics* (Gigerenzer et al,1996) which do not employ utilities in decision making. Instead, fast and frugal heuristics guide decision making through the use of simple search and stopping rules {Lloyd 2003:398}. Gigerenzer et al, (1996) suggests that respondents compare one aspect at a time, if they are able to distinguish between two choices then a decision is made on that basis and all other information is ignored. This approach is often referred to as '*Take the Best*' {Gigerenzer et al, (1996); Lloyd 2003}.

#### *Disease label*

When discussing the PTO method respondents again tended to refer to the disease label as their main focus.

*"I think I was probably focussing more on the disease label in this, rather than the breakdown here (points to functional health status).I think maybe, perhaps, quadriplegia, because it is quite horrible, having to sort of think, that all these people are supposed to die, I probably disassociate from it a bit. And that's why I think I stuck just to the disease labels" (ID 19: panel 3).*

The reference to the disease label and its importance in the decision making process was explored further in the interviews. Some respondents referred to the additional complexity the



disease label introduced to the PTO exercise. For example, there were suggestions that the disease label personalised the people and increased the difficulty respondents face when taking the unfamiliar role of decision maker in the health context.

*“Having the disease names made it more difficult in a way, because they’ve got like, kind of similar in terms of, I don’t know, unable to perform the usual activities, but I was trying to get rid of the stigma of quadriplegia as like one of the worst things you could have, and then severe depression, and then my conscience plays into me as well, because like I am thinking, ok, so how about all these people with severe depression, oh well, all those people with quadriplegia, it’s awful. Sorry” (ID 2: panel 1)*

Again it would seem that respondents are adapting their decision making behaviour and focusing on certain attributes more than others, with particular reference to quadriplegia, which was perceived to be the ‘worst’ health state by the majority of respondents.

#### *Functional health status*

When discussing the TTO exercise respondents tended to focus on the disease label when referring to the more severe health states. When the choice between health states was less clear, i.e. the disease states were similar and generally thought to be less severe, respondents would often look at the number of dots on the functional health states.

*“It was different for TTO I sort of just looked at the name, you know quadriplegia, and decided ..but if they were similar I don’t know like Asthma and Stroke I would use the dots to help me decide” (ID 3:panel 1)*

Again this respondent seems to be adapting the *take the best* heuristic, by taking the most important attribute first (i.e. disease label) and if it was possible to distinguish between choices then a decision could be made at that point. If it was not possible to distinguish between the choices then the respondent would move on to consider their second most important attribute (i.e. dots on the functional health status).

The use of the dots on the functional health status to help inform respondent’s decisions to the various exercises was a recurring theme. The dots were used to describe the level of severity of each attribute (no dot equals no pain, one dot equals some/ moderate problems and two dots equals extreme problems).

*“I used the dots a lot. Because a lot of them were similar anyway. Like the coronary heart disease one and the depression. Initially you think, oh coronary heart disease, and then I used the dots a lot to kind of rank them... I mean I had a quick read through the*

*description but I really focussed everything on the dots and how it was affecting them”  
(ID 10: panel 2)*

As mentioned above when making choices for diseases that were seen as less severe than say quadriplegia other cues such as the dots on the functional health status became the focus. The order in which cues are used is dependent on what Lloyd refers to as their cue validity.

Cue validity is determined by the success or not of past experience. In the panel sessions the TTO exercise was administered last so one would expect respondents to have had some learning experience and to have familiarised themselves with the health states during the previous exercises (ranking, VAS and PTO). One interpretation is that the whilst the cues used in the TTO exercise differ from the other exercises respondents are still using fast and frugal heuristics.

*“The TTO was much easier than the other one, but I new the diseases and what I was expected to do by then. I remember thinking how did I rank, or what did I think last time and then made my decision based on that....like oh say quadriplegia, well I would hate to be in that state so I will give up more time, then say breast cancer not that bad ok I will not give much time for that”(ID 19: panel 3).*

*“I sort of new what I was doing by then and we had really thought and discussed the cards in lots of detail, so I sort of new what they were about I just needed to get my head around what the TTO question was asking...which was helpful because there was not that much to think about really, just the general question” (ID 6: panel 1).*

During the interviews it became apparent that certain dimensions on the functional health state descriptions were more influential than others in informing decisions. Dimensions that were more frequently mentioned were anxiety and depression, and cognitive functioning.

*“I would be able to cope better with physical problems than with mental ones. So, I was looking at anxiety and depression and cognitive functioning. So, the physical side not so much of a problem I tended to look at the mental stuff” (ID 5: panel 1).*

The main question around the stimulus is did respondents ignore the other attributes – supporting Gigerenzer et als (1996) theory of fast and frugal heuristics or did the engage in compensatory decision making. If the latter then they would be trading between attributes, rather than ignoring attributes, and therefore valuing one attribute higher than another. For example, in the above quote more value would be placed on the attributes of anxiety and depression and cognitive functioning.

*Undertaking the VAS exercise*

One of the key advantages of the VAS is that respondents tend to find the exercise relatively easy (Brazier et al (1999); Parkin and Devlin (2006)). The respondents in this study also tended to suggest that the VAS exercise was easy to undertake. When explored further with respondents they tended to refer to the simplicity of the exercise in relation to the content of the decision task.

*“Once you had thought about the diseases and ranked them then what you are really being asked to do is not that difficult really, it’s just an extension of the ranking exercise” (ID 19: panel 3).*

When respondents were asked to explain how they placed the 15 diseases onto the VAS, the majority of respondents mentioned ranking or placing cards in a sort of order before placing onto the scale.

*“I just went through the cards, reading what was on there and deciding what living in that state would be like, I mean something like breast cancer and then I put them in a sort of order that I imagined would maximise health and then went through and put down scores” (ID 10: panel 2)*

This suggests that respondents were able to undertake the exercise as it was intended, in so far as they were considering all the information and were seeking to maximise utility.

A recurring theme was the use of end points, with the extreme health states being the first and easiest to place on the scale. Respondents suggested that they used the extreme states to judge the other diseases against. It was also implied by respondents that the diseases that were placed at the end of the scale were most likely to be given a ‘stable’ value, whereas, the diseases which were placed towards the middle of the scale may have been more subject to change. Thus, it would seem that respondents had more confidence in the reliability of their extreme values.

*“ I felt happy with my best and worst health states and these are the least likely to change however, there were a group in the middle that may move around each time I did this yeah in the end I just went for it” (ID 5: panel 1).*

A common theme was the grouping of diseases. The quote below describes how one respondent actually categorised disease states prior to placing them on the VAS.

*“So you could say I’d given diseases like 5 classes, like class A, class B, class C, class D, class E and all class A was similar, similar area with slightly higher numbers for the ones that were worse. I did this first then thought about them in terms of the scale thing” (ID 3: panel 1).*

*What do the numbers mean?*

There is debate in the literature around whether the VAS actually measures strength of preference or is just merely a ranking exercise with respondents treating the numbers on the scale rather arbitrary. In the panel session respondents were asked to think about the distance between the numbers. The interviews explored this with respondents.

There were a mix of responses with some respondents suggesting that they did use the numbers and the distance between the numbers did mean something.

*“Yes the distance between the numbers was important – I remember thinking well if that is if I put a disease at say 34 and one disease at 44 there was a distance of 10 points between these diseases” (ID 20: panel 3)*

Whilst others found it difficult to consider the numbers when undertaking the VAS exercise.

*“And in a way, you know, because the distance between them was more important, sometimes it, say, if you put something at 54, then that meant that if you thought that another disease was kind of four spaces away worst than that, then you'd have to put something there. So it could get quite crowded, and then I could not decide and you realise you were getting in a bit of a mess, and then to do it again. Then I think I just stopped worrying about the numbers and put them in order of importance. I used most of the scale” (ID9: panel 1)*

In contrast to the above quotes, other respondents suggested they did not attempt to use the numbers on the VAS.

*“I remember making this conscious decision that I was going to use the whole scale and not be stuck in that kind of, that very small spectrum, I think you know, after that I'm not quite sure why I would have said something was 68 rather than 84” (ID 17: panel 2)*

*Interviewer : “what was your reason for using most of the scale?”*

*“Because I.. well if I am honest, because I could. I suppose my worst and best were the extremes you know near top and bottom, then after that I just filled the gap” (ID 4: panel 1)*

The interviews suggest that some respondents followed the instructions outlined in the panel session and did seem to use the numbers to measure their strength of preference, whilst other respondents did not use the numbers and implied they simply used the scale to rank the disease states.

*Undertaking the PTO exercise*

Most respondents suggested they had difficulty undertaking the PTO exercise. Stating that they found it more difficult than both the VAS and TTO exercises. When this was explored further during interviews respondents said that this was partly due to the framing of the question. Respondents suggested that they had difficulty 'grasping' what was expected of them.

One aspect was the amount of information presented to respondents.

*"I think there were just so many kind of other, you know, with a hundred people, they're going to die within between two and four years, so you've already got some variation. You know, is it two or is it four, or somewhere in between. Then you've got, you don't know who these people are, well is that significant or isn't it? So you're kind of thinking about that. The assumptions brought something else in for you then as well, because they brought more things to think about rather than taking things out. And trying to work out, well what is it that you need to keep in your head here. Because if, if all of these things are equally important, then every time you make a judgement you have to go back and think about all of these things"(ID 1: panel 1).*

As mentioned in Chapter 6 quadriplegia was used as the anchor state (i.e. all health states were measured against it). A recurring theme was the use of the use quadriplegia as the anchor state which made the task problematic.

*"I struggled with the wording of the question, it took me ages to get it into my head what was being asked. But I think in the end, I kind of worked it through. Quadriplegia does play a big part. I just want to treat people with quadriplegia" (ID 2: panel 1*

One respondent referred to 'getting hung up' on quadriplegia.

*"I sort of got hung up on quadriplegia and could never get passed it, you know, I mean nothing could be worse than that...so quadriplegia won every time for me I would also treat those people"(ID10: panel 2)*

Again this behaviour may be explained by Gigerenzer's 'Take the Best' approach with the respondent making the decision based on one aspect of the task, in this case the anchor state (quadriplegia) and generally ignoring other information.

A further theme which emerged when discussing why respondents found the PTO exercise difficult was around contextual issues. The PTO differs from the other two methods in so far as it takes a societal rather than an individual perspective. Most respondents commented on the difficulty of thinking and making decisions for others rather than themselves.

*“When it’s for you its ok, but having to make a decision on whether other people should live or die no that was difficult I really could not get my head around that” (ID1: panel 1)*

Respondents suggested they had some difficulty understanding the PTO question and what they were being asked to do. Respondents seemed to adopt various strategies to answer the questions. For example, one respondent refers to using their own knowledge of prevalence and statistics as a way to answer the PTO questions.

*“I tried to think, although I haven’t got a knowledge of prevalence of diseases and statistics, but that’s what I tried to think about. I just think, well how many people, sort of would have this disease, so how many lives are we saving? It probably wasn’t that accurate, because I didn’t have that information to hand, and I don’t know, But your own sort of analysis, you might have some knowledge, media knowledge coming in there, or what you’ve read in the paper” (ID 5: panel 1)*

Respondents also had a problem with the hypothetical context of the question and the abstract nature of the exercise.

*“I think there was just a lot going on in that, and again it goes back, it’s very difficult to kind of do this as an abstract exercise, without other things crowding in. Because of course the identity of these people is going to be important, and where they are, and all those other things. So, just kind of actively trying to see this as a sort of abstract, could be anywhere, could be anytime, I just found it really, really difficult, and that I found quite frustrating.... it’s out of context, so you have no sense of the consequences of your actions. And so, because it’s not a real situation, it is quite difficult to think, well, are these the only factors that would influence me. Of course they’re not” (ID 19: panel 3).*

A common statement throughout the interviews was the reference to ‘playing the numbers game’ this is evident in the statement below.

*“Thinking about others is difficult, but then, I knew it wasn’t for long, so one valuation and it is not real. It is just like playing a number crunching game”(ID 14: panel 2).*

#### *Undertaking the TTO exercise*

Respondents tended to suggest that the TTO exercise was easy to moderate although, respondents did comment on the hypothetical nature of the task.

*“Yes it was ok, a bit unreal not true to life but not hard to understand” (ID 19: panel 3)*

One reason for the more positive response was that they tended to compare the exercise with the PTO. Thus, the response may have been different if only VAS and TTO exercises had been administered.

*“I think I found it easy because PTO was so difficult for me” (ID2: panel 1)*

It was suggested that the context of the TTO question and the fact that respondents were being asked to think about themselves rather than others made the exercise easier.

*“Yes, again, I think it's because I could more relate it as a personal decision rather than thinking of others.” (ID 1: panel 1)*

*“It was easier to focus on what it would be like for an individual than it would for 50,000 or something” (ID 6: panel 1)*

A further theme that helps to explain why some respondents found the TTO simpler than the PTO was due to the fact that discussion around aspects, such as, the health states etc had been aired earlier in the session, and respondents were more accepting of what they were being asked to do. Another explanation offered by one respondent was that she was tired after four and half hours valuing health states so she just did what was asked.

*“I think I did just disregard that yes, it had been a long day and I knew what you wanted” (ID 17: panel 2)*

#### *Trading life years in TTO*

All respondents said they traded some life years for certain diseases.

*Interviewer “So you found yourself trading life years for some diseases and not others?*

*“Yes, there is a threshold in these the way you, er, there are certain conditions like the worst imaginable health state the you would give up all your life for a week or whatever rather than be in there you would rather be dead or something, that ones set. Then there is another whole batch probably anything with one dot or may be with two dots which is, unless it is on pain or cognitive functioning, were I mean realistically there is no such thing as a best imaginable health state nobody is in that so you don't live now in the best imaginable health state the state you are living in now could easily be one of these one dot things. So, that's your choice” (ID12: panel 2)*

All respondents completed the exercise, however, when asked why they are willing to trade life years, comments from most respondents suggested they were not taking this aspect of the session very seriously. A theme that was evident from a number of interviews was that respondents

traded life years because that was what they were expected to do. Respondents again referred to 'playing the game.'

*"I did, but I'm not quite sure how serious I was about it. I was a bit ambivalent about it towards the end. By that time I think I was just playing the game" (ID 19: panel 3).*

### *The numbers*

As discussed in Chapter 5 the validity of the methods may be dependent on the numeracy of respondents. With lower numeracy interfering with the completion of cognitively complex exercises such as TTO and PTO (Woloshin et al 2001: 386) add the other ref around difficulty with probabilities and simple maths). Something which became evident when undertaking the range reflection in the panel session, and a point that was recorded in the observer notes, was the difficulty some respondents had in terms of working out the implications of their choices. For example, if someone was to choose 50,000 people before they would choose programme B the facilitator would ask the respondent the following "do you mean that you need 500 people in programme B for every one person in programme A."

Some respondents highlighted their lack of confidence in using numbers and how this influenced how they did the exercise.

*"I don't work with numbers it is not the kind of thing I am comfortable with, I mean, yes it means thinking about the differences, if your not used to numbers, you don't necessarily think that there is a difference, there is a four point difference between diseases. I think the numbers did have an impact, because if people aren't comfortable with the it does have an influence on how they do the exercise and the kind of results they get" (ID 4: panel 1)*

For the PTO exercise respondents seemed to have problems with the use of large numbers. Again being able to visualise things in this case people was an important part of the valuation process for many.

*"I did find it difficult, because the numbers, I think the numbers, and I can't sort of visualise, I'm trying to think of the example you gave and I still, I just could not comprehend" (ID 10: panel 2)*

There was discussion around how and indeed if respondents managed to overcome this number problem. Respondents again referred to using visual images to inform their response to the PTO questions.



*“I mean 100 people with a mild condition and 1 person for quad you would probably go for quad. But if 1000 people in a mild condition starts to get quite sizeable. I mean that was like thinking about half a school or something. I did think like how big is a school, or how big is Villa Park or Wembley stadium or something like that. I tried to visualise the whole of Villa Park with a disease. Once you get to Wembley stadium it gets a bit arbitrary” (ID 19: panel 3)*

#### *Use of responses in decision making*

Finally, the interviews explored respondent's thoughts on the use of their values elicited in the panel process to inform healthcare decision making.

There was a mixed response to this question with some respondents feeling confident about their values being used in a decision making context:

*“Yes why not my values are as good as anyone else's...so I would have no problem with them being used” (ID 19: panel 3).*

In contrast, other respondents state that they would not like their responses used in decision making:

*“No I would not I don't feel qualified to make those kind of decisions, or experienced enough to be honest...I mean to be honest, I don't think I could ever be really happy with those, because I can't take out the person, the personal, human aspects of it. I don't, like these you can kind of say, oh yes I can trade my own life because it's dealing with me. But dealing with other people, I can't dehumanise it and say, I'll give this much money to these people and not these, without thinking of the people involved, while these were, it's my own life, it's my own decision. It just made you think really, just things you never really think of. I'm glad you don't have to think about that, well not glad you don't have to think about it, glad you don't have to make the decisions” (ID2: panel 1).*

#### **Discussion**

Previous studies have tended to use quantitative analysis to explore aspects of validity and reliability of health state valuation techniques. There are an increasing number of academic publications that conduct quantitative analysis and call for more qualitative research (Dolan, 2000).

The respondents who took part in the interviews are 'informed respondents' in that they are discussing health state valuation exercises they have previously completed. The gap between the health state valuation exercise and the interview is advantageous in so far as it allows people time to sufficiently reflect on exercises. Such time allows respondents to reflect on topics for discussion and to enable individuals to shift their cognitive framework from the acceptance of a

series of techniques required to allow people to engage in the valuation exercises. It would also not have been appropriate to undertake interviews immediately following a five hours preference elicitation exercise. However, other methods such as 'talk aloud' or interviews immediately following panel sessions may capture maximum participation as well as the exercises being fresh in people's minds.

The study further reaffirmed the theory that valuing hypothetical health states is a complex task. Most respondents found the exercises difficult, having said that, all respondents said that they did trade for some, if not all diseases, when undertaking the PTO and TTO exercises. The interviews explored a number of areas that are relevant to assessing the validity of the instrument and process of elicitation.

Economic theory predicts that individuals are utility maximisers. In this context, it suggests that decision makers are able to use available information in a logical and systematic manner, in order to maximise their utility. In contrast, cognitive psychologists 'consider an interactive process where several factors may influence a decision in a non-trivial way' Kahneman et al (2002). Psychologists such as Kahneman et al (2002) also suggest that a number of factors including emotions, attitudes, environment and past experience may all effect the decision making process with behaviour being adaptive and dependent on the context and transitory perceptual conditions. In this study the focus on the disease label and the knowledge of conditions such as quadriplegia seemed to support this view.

The psychological literature also suggests that individuals have difficulty in fully analysing situations that involve economic and probabilistic judgements (see Kahneman et al 2002) and the more complex the task the more difficult this becomes for individuals. Therefore, individuals rely on heuristics or cognitive shortcuts when can be biased i.e. individuals may ignore relevant and often important information when making their choices. This study goes some way to support this approach with evidence that some respondents were using non-compensatory decision making, with individuals often using one or two attributes when making their decisions.

It did seem that some respondents may have been adopting the 'take the best' heuristic. The question here is if individuals adopt such a strategy are they simplifying the task i.e. is this a 'smart heuristic or are they simply ignoring other factors and taking a 'second best option'. Existing exploratory work around these heuristics tends to look at inferences rather than preferences. For example, studies have asked respondents which is the largest city, respondents

use the recognition heuristics i.e. they chose a city they have heard of. When looking at inferences there is generally a gold standard against which the decision heuristic can be compared to assess whether it is smart or in fact a limited approach to decision making. This is not the same for preference elicitation when the gold standard is maximising outcome (Ryan and San Miguel 2003).

There did seem to be some relationship between complexity of task and the use of fast and frugal heuristics for example undertaking VAS was generally reported as easy with many respondents suggesting they used all of the information presented. Whilst for PTO and TTO they tended to focus on one or two attributes, which for PTO tended to be the anchor state quadriplegia.

In this study if respondents were using non-compensatory decision strategies questions arise around the frequency i.e. were they consistently using the same simplifying approaches or did this change between methods and or stimuli presented. Again raising further questions around whether respondents were maximising utility.

The interview data suggested that respondents had more difficulty with the societal perspective and the task of trading people in the PTO, than they did with the individual task of trading time in the TTO exercise. This supports the empirical work conducted by Damschroder et al 2005 who suggest that 'within the context of a PTO elicitation, people are more likely to consider non-maximising principles'. Such findings question the validity of the PTO technique in the QALY setting.

The stimuli presented are an important factor in the valuation process. In this study both a disease label and functional health status were used. Unlike generic stimuli the disease label contains information of the underlying disease. This information may be essential, or introduce undesired effects, or both. A study conducted by Robinson and Bryan (2001) showed that quadriplegia had a negative effect on valuations, similar results were found when the term 'cancer' was presented Gerard *et al* (1993) and Rabin *et al* (1993). In this study the disease quadriplegia seemed to have a negative impact on the valuation exercises for some respondents. This could be due to both the severity of the disease and the knowledge or 'baggage' individuals already have in terms of different diseases.

So where does this leave us? Evidence from this exploratory study challenges the economic theory that underpins the QALY approach. Having said that expected utility theory is simply a model that helps us to explore the complex world of decision making- one could compare this to

the perfect market model which allows us to explore and measure activity in economies. Such models are not reality but conceptual short hands that help us to make sense of the complexity of the real world. In this case explore the nature of decision making in the area of health state valuation. All models have limitations, and it is important that we explore and acknowledge such limitations. In terms of preference elicitation we need to further understand the complexity involved in decision making tasks and not only explore the outcomes but also the process - gaining a greater understanding of how and why respondents act in the way they do.

Therefore the challenge for health economics is to look beyond economics and use the work of other disciplines such as psychology and sociology to help in the refinement of health state valuation techniques and the area of preference elicitation.

## References

- Berg, B.L. (1998). *Qualitative research methods for the Social Sciences* (3<sup>rd</sup> ed). Boston: Allyn & Bacon.
- Brazier J, Deverill M, Green C, Harper R, and Booth A. (1999) A review of the use of health status measures in economic evaluations. *Health Technology Assessment*;3(9).
- Cairns, J., van der Pol, M., Lloyd, A, (2002). Decision making heuristics and the elicitation of preferences: being fast and frugal about the future, *Health Economics*, 11(7)655-658
- Damschroder, L.J., Roberts, T.R., Goldstein, C.C., Miklosovic, M.E., Ubel, P.A., (2005). Trading people versus trading time: what is the difference? *Population Health Metrics* 3:10
- Dolan, P., Gudex, C., Kind, P. (1996). Valuing health states: a comparison of methods, *Journal of Health Economics*, 15, 209-231
- Dolan, P. (2000). "The measurement of health-related quality of life for use in resource allocation decisions in health care," Chapter 33, pages 1761-1799 in Culyer, A.J. & Newhouse, J.P. (ed), *Handbook of Health Economics*, edition 1, volume 1
- Gerard K, Dobson M, and Hall K. (1993). Framing and labelling effects in health descriptions: Quality adjusted life years for treatment of breast cancer. *Journal of Clinical Epidemiology*; 46:77.
- Gigerenzer, G. et al (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, 103(3): 592-596
- Kahneman, D., Solvic, P., Tversky, A., (2002). *Judgement under uncertainty, heuristics and biases*, Cambridge. CUP
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*, London, Sage
- Mallinson, S. (2002) Listening to respondents: a qualitative assessment of the short-form 36 health status questionnaire, *Soc Sci Med*, 54(1):11-21.
- Lloyd, A. (2003) Threats to the estimation of benefit: are preference elicitation methods accurate? *Health Economics*, 12(5):393-402
- Loomes, G. (1998) "probabilities vs money: A test of some fundamental assumptions about rational decision making", *Economic Journal*, 108:477-489

- Marshall, C., Rossmans, G.B.(1999). *Designing qualitative research*, (3<sup>rd</sup> ed), London, Sage Publications.
- Mcguire, A., Drummond M. (eds) (2001).*The theory and practice of economic evaluation in health care*. Oxford, OUP
- Miles, M., Humberman, M. (1994) *Qualitative data analysis: an expanded sourcebook* (2<sup>nd</sup> ed), London, Sage Publications
- Parkin, D., Devlin, N. (2006) Is there a case for using visual analogue scale valuations in cost-utility analysis, *Health Economics*, 15:653-644
- Patton, M. (1987). *How to use qualitative methods in evaluation*, London, Sage Publications
- Payne, J.W., Bettman, J.R., Johnson, E.J (1993). *The adaptive decision maker*, Cambridge, University Press
- Pope, N., Mays, N. (2000). *Qualitative methods in health care* (2<sup>nd</sup> ed), London, BMJ
- Rabin R, Rosser RM, and Butler C. (1993) Impact of diagnosis on utilities assigned to states of illness. *Journal of the Royal Society of Medicine*;86:444.
- Read, J.L., Quinn, R.L., Berrick, D.M et al, (1984). Preferences for health outcomes comparison of assessment methods, *Medical Decision Making*, 4(3), 315-329
- Robinson and Bryan (2001) 'Naming and framing': an investigation of the effect of disease labels on health state valuations, Paper presented at the Health Economists' Study Group Meeting, January 2001, University of Oxford
- Ryan, M., San Miguel, F. (2003). Revisiting the axiom of completeness in health care, *Health Economics*, 12(4):295-230
- Silverman, D. (2000). *Doing qualitative research: A practical handbook*, London, Sage
- Torrance, G.W. (1976). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5:1-30
- Woloshin, S., Schwartz, L.M., Moncur, M., Gabriel, S., Tosteson A.N. (2001). Assessing values for health: numeracy matters, *Medical Decision Making*:21(5):382-390

## **Appendix 1**

### **Planned health state valuation exercise**

The sessions had six main stages these included introduction to the health states, a warm up exercise and the three different valuation tasks.

#### *Stage 1: introduction to health states*

During stage one the facilitator went through the health states highlighting the disease label and explaining the functional health status, in terms of the different dimensions and levels of severity. There was no discussion of the health states at this point but participants could ask points of clarification.

#### *Stage 2: warm up exercise*

Stage two consisted of a warm up exercise where participants were asked to rank the health states. This was done individually and did not involve discussion with the wider group.

#### *Stage 3: VAS exercise*

For stage three of the panel process participants were asked to place the health states on the VAS, again they did this individually without discussion with the group.

#### *Stage 4: PTO exercise*

For stage four of the panel process participants were asked to undertake the PTO exercise.

PTO Format:

*Imagine that you are a decision maker. You have a choice between two programs that will reduce the incidence of disease in a few years from now.*

*Program A will prevent the occurrence of a rapidly fatal disease in 100 people in your country.*

*Program B will prevent the occurrence of disease X (chronic state described in detail) in N people in your country.*

*The programs are in all other respects equal.*

*Choose the value for N that would make you indifferent between the two programs. When answering, please disregard possible economic aspects.*

The exercise began with the facilitator introducing the PTO method using an example of severe vision disorder. Participants were then asked to individually value the 9 chronic diseases using the PTO method. Respondents were then given the opportunity to discuss their responses with the group and finally participants had a chance to change valuations.

#### *Stage 5: TTO exercise*

For stage five of the panel process participants were asked to undertake the TTO Method. The same process that was used for the PTO method was used i.e. the facilitator explained the TTO method using the example of severe vision disorder, participants were asked to undertake the process individually, followed by group discussion and finally participants had a chance to change valuations.

TTO format:

**Consider living for 10 years in the health state described above**

The alternative option is living in full health for 10 years or less

Ask yourself which point you are no longer able to choose between the two options: your indifference point

- Consider whether you are prepared to trade off life years to avoid getting the health state described above and be in full health?  
**YES / NO**
- I am indifferent between 10 years with the health state described above and  
**Years in full health.**

*Stage 6: final opportunity to change responses*

Finally, participants were asked to consider all their individual valuations for the VAS, PTO and TTO and were given a final opportunity to alter any of their responses, there was no discussion at this point.

It was expected that panels would last 5-6 hours. This process was designed to allow for discussion and deliberation and participants had the opportunity to change their PTO and TTO responses on 3 different occasions during the panel session.