

# Department of Economics

Are Some Deaths Worse Than Others?

The Effect of 'Labelling' on People's Perceptions

Anne Spencer, Judith Covey, Angela Robinson and Graham Loomes

Working Paper No. 620

December 2007

ISSN 1473-0278



Queen Mary  
University of London

# Are some deaths worse than others? The effect of ‘labelling’ on people’s perceptions

Anne Spencer<sup>1</sup>, Judith Covey<sup>2</sup>, Angela Robinson<sup>3</sup>, Graham Loomes<sup>3</sup>

<sup>1</sup> Queen Mary, University of London

<sup>2</sup> University of Durham

<sup>3</sup> University of East Anglia

## Abstract

This paper sets out to explore the extent to which perceptions regarding the ‘badness’ of different types of deaths differ according to how those deaths are ‘labelled’ in the elicitation procedure. In particular, we are interested in whether responses to ‘contextual’ questions – where the specific context in which the deaths occur is known – differ from ‘generic’ questions – where the context is unknown. Further, we set out to test whether sensitivity to the numbers of deaths differs across the ‘generic’ and ‘contextual’ versions of the questions. We uncover evidence to suggest that both the perceived ‘badness’ of different types of deaths and sensitivity to the numbers of deaths may differ according to whether ‘generic’ or ‘contextual’ descriptions are used.

Qualitative data suggested two reasons why responses to ‘generic’ and ‘contextual’ questions differed: firstly, some influential variables were omitted from the ‘generic’ descriptions and secondly, certain variables were interpreted somewhat differently once the context had been identified. The implications of our findings for ‘generic’ questions, such as those commonly used in health economics (for example, the EQ 5D), are discussed.

**J.E.L. classification codes:** H5, I10

**Key words:** preferences; context effects; affect heuristic.

## **BACKGROUND**

People regard some deaths as worse than others.<sup>(1)</sup> For example, research has shown people are more concerned about cancer deaths than deaths from heart disease, motor vehicle accidents, household fires, or airplane accidents.<sup>(2,3)</sup> Likewise deaths caused by industrial air pollution are regarded as more deserving of resources than deaths caused by smoking or automobile accidents.<sup>(4)</sup> Hence, it would appear that just as the acceptability of risks can be characterized by “qualitative” factors,<sup>(5)</sup> public concern about deaths might be aggravated by certain underlying features of those deaths.

As noted by Sunstein<sup>(1)</sup> the risk perception literature does suggest reasons why people might regard some deaths as worse than others. More specifically, as well as the notion of livable life-years (i.e., it is worse if a child is killed than an older adult), he highlights the importance of ‘dread’ (i.e., death preceded by unusual pain and suffering), blameworthiness (i.e., responsibility for death lies with a third party), distributional equity (i.e., victims are members of socially disadvantaged groups), and high externalities (i.e., catastrophic events involving widespread non-pecuniary losses). For example, cancer deaths might be more “dreaded” than deaths from heart disease because they are preceded by a longer period of pain and suffering, whereas people who die from industrial air pollution are less to blame than smokers are for their deaths.

If people’s concerns about deaths can be captured by a set of underlying attributes such as these the next step would be to quantify their importance. This type of quantification would then allow more general conclusions to be drawn about the ‘bad deaths’ premium that people place on different types of deaths according to their underlying features rather than to their specific identity. However, before attempting such a quantification we need to make sure that

people's concerns for different types of deaths can be adequately captured by the 'generic' attributes used to describe the deaths. The research published in this paper explored this issue for deaths that were described using four of Sunstein's attributes – the age of the victim (livable life-years), the severity and duration of the victim's pain and suffering in the period leading up to their death (dread), and who is most to blame for the victim's death (blameworthiness). More specifically, the aim of the study was to test how well people's concerns towards these attributes only – 'generic' descriptions of deaths – matched with their concerns towards descriptions where the specific cause of death is also identified (for example, as a driver in a car accident) – 'contextual' descriptions.

Literature from a related area, health state utility measurement, gives us reason to suspect that the revelation of the identity or 'labelling' of the cause of death might make a difference. For example, Sackett and Torrance<sup>(6)</sup> found that adding a label to a health state description significantly affected the utility values obtained: 'tuberculosis' was given a higher value than 'unnamed contagious disease', whereas 'mastectomy for breast cancer' was given a lower value than 'mastectomy for injury'. Gerard et al.<sup>(7)</sup> also found differences when the word 'cancer' was used and when descriptions were written in the third party. Likewise, Rabin et al.<sup>(8)</sup> and Robinson and Bryan<sup>(9)</sup> report that adding a label significantly affected valuations of both physical and mental conditions, but found differences in the direction of these effects.

Of course, evidence that responses to 'generic' and 'contextual' descriptions are *different*, does not, in itself, imply that one description is somehow superior to the other. Before reaching such a conclusion, as well as considering the purpose to which the responses are to be put (we return to this issue in the discussion), we need to establish which set of responses is the more valid. One objective criterion against which to assess the validity of responses is

to test the sensitivity to a theoretically relevant factor that should influence people's concern such as the *number of deaths* caused by a particular type of death (see Loomes<sup>(10)</sup>). Sensitivity to the numbers dying, therefore, offers a test of the validity of responses to both 'generic' and 'contextual' descriptions.

To sum up, the aim of our study was to explore the extent to which perceptions regarding the 'badness' of different types of deaths differ according to whether 'generic' or 'contextual' descriptions were used, where deaths are 'labelled' in the 'contextual' descriptions. Further, we set out to test whether sensitivity to the numbers of deaths differed across the 'generic' and 'contextual' versions of the questions.

## **METHODS**

### **Sample**

Discussion groups of between 8 and 12 participants were convened in the North East and East Anglia regions of the UK. A total of 313 participants were recruited to take part on a quota basis to be broadly representative of the gender, age and social class profile of the general population. The demographic characteristics of the sample are shown in Table 1.

TABLE 1 ABOUT HERE

### **Group Meetings**

The group discussions began with a brief introduction to the aims of the study and participants were asked to answer in their capacity as citizens. It was emphasized that the research would provide guidance to the Health and Safety Executive (HSE) about whether

members of the public wanted more effort and resources to be put towards preventing some sorts of deaths than others. More specifically, they were told that the HSE were interested in finding out whether they wanted to take into account factors like the age of the victims, how much and for how long their quality of life is affected in the period leading up to their deaths, or who is primarily to blame for the deaths.

Following the introduction participants were presented with a series of questions, each of which presented them with two premature death scenarios, labelled A and B, which varied on one or more of the five attributes shown below:

- The **numbers** of people who die (either 10, 15, 25, or 50)
- Their typical **age** (under 17s, 17-40s, 40-60s, or over 60s)
- How **much** their quality of life is affected in the period leading up to their deaths (either a bit worse than normal or a lot worse than normal)
- How **long** their quality of life is affected in the period leading up to their deaths (either a few minutes, a couple of weeks, 1-2 years, or 3-5 years)
- Who is most to **blame** for the deaths (the individuals themselves, other individuals, business/government or nobody in particular)

The two premature death scenarios used in each question were presented in tabular form as shown in the example presented in Figure 1. In each question participants were asked to rate which of the two types of deaths they thought was the worst by ticking one of four possible responses, namely: A is much worse than B; A is slightly worse than B; B is slightly worse than A; or B is much worse than A.

FIGURE 1 ABOUT HERE

Participants were presented with a total of 28 questions like this. Five of these questions were practice questions in which one attribute was varied at a time and 12 questions made up a discrete-choice experiment that was used to estimate the weights of each attribute – the details of which is reported elsewhere (Robinson et al.<sup>(11)</sup>). We report here the 11 questions that set out to explore the impact on responses of providing participants with information about the specific causes of the deaths over and above information about the deaths in terms of the five attributes described above. Six of these questions involved scenarios where the specific causes of death were not provided – we call these the ‘generic’ questions. These six ‘generic’ questions were presented along with the 12 questions which were used for the discrete choice experiment in Part 1 of the questionnaire. Two of the six ‘generic’ questions were identical to one another and were presented as the 6<sup>th</sup> and 21<sup>st</sup> questions, providing a test-retest reliability check on responses to ‘generic’ questions. In Part 2 of the questionnaire the five different ‘generic’ scenarios were presented again but this time the specific causes of death were also presented – we call these the ‘contextual’ questions. For example the causes of death provided for the ‘generic’ question shown in Figure 1 were ‘deaths from lung cancer caused by smoking’ for A and ‘deaths from asbestos-related cancer’ for B.

Three versions of the questionnaire were produced each of which used different types of deaths for the five ‘generic’/‘contextual’ question pairings. The ‘contextual’ questions used in each version are shown in Figure 2.

FIGURE 2 ABOUT HERE

Between the three versions eight different causes of deaths were used to reflect the range of variation on the attributes (i.e., car drivers, pedestrians, rail passengers, cancer caused by

smoking, cancer caused by asbestos, cancer caused in the workplace, accidents at work, breast cancer and carbon monoxide (CO) poisoning). These causes of deaths were paired up in a total of seven different ways (car drivers vs. rail passengers, car drivers vs. pedestrians, pedestrians vs. breast cancer, work-related cancer vs. car drivers, accidents at work vs. car drivers, CO poisoning vs. accidents at work, and smoking cancer vs. asbestos cancer).

All the pairings were included in two of the three versions apart from car drivers vs. rail passengers which was included in all three versions and used in its 'generic' form as the repeated question for the test-retest reliability check. However, as shown in Figure 2 the level of the number of deaths attribute was different for each pairing between the versions.

These questions can be thought of as falling into one of two categories. In some cases, evidence from piloting suggested which way a weighting would go, and the objective was to get some indication of the strength of that weighting. For example, we knew from piloting that deaths from asbestos-related cancer would be given more weight than deaths from lung cancer attributed to the individuals themselves being smokers. So the numbers in both the smoking cancer vs asbestos cancer comparisons were increased in the same direction, in one case in the ratio 1.5:1 and in the other case in the ratio 2:1; likewise for work-related cancers vs car drivers and for accidents at work vs car drivers. This variation therefore enabled us to test whether respondents' ratings were equally sensitive to the different ratios of deaths in the 'generic' and 'contextual' versions of each pairing.

However, for pedestrians vs breast cancer, CO poisoning vs accidents at work and pedestrians vs car drivers the numbers were counterbalanced (i.e. appeared in the ratios of 1:1.6 and 1.6:1) in each direction because we had no strong priors.



Respondents had not been asked to explain their decisions for the ‘generic’ questions, but for each of the ‘contextual’ questions a box was provided at the bottom of the page and respondents were invited to write a sentence or two giving their reasons for their answers.

In summary, the questionnaire contained the following four main parts:

- 1) A test-retest reliability check.
- 2) A test of the differences between the ‘generic’ and ‘contextual’ responses.
- 3) A test of the sensitivity to different ratios of deaths in the ‘generic’ and ‘contextual’ responses.
- 4) Thematic qualitative analysis of the written reasons given to the ‘contextual’ questions.

The test-retest reliability check was designed to check that any observed differences were due to a genuine difference between the ‘generic’ and ‘contextual’ responses, and not an artifact of the questions being asked twice. The second set of tests were designed to check the degree of any differences between the ‘generic’ and ‘contextual’ responses. The third set of tests were designed to check if the ratings were equally sensitive to the different ratios of deaths in the ‘generic’ and ‘contextual’ versions of each pairing, and allowed us to consider if the ‘contextual’ question led to different weights on the attributes presented in each pair. Finally the thematic qualitative analysis allowed us to examine why participants’ preferences might have been different between the ‘generic’ and ‘contextual’ versions of each pairing.

## **Analysis**

In the ‘generic’ and ‘contextual’ questions the four response categories were scored 1-4, i.e., 1= A is much worse than B, 2= A is slightly worse than B, 3=B is slightly worse than A, 4=B is much worse than A. The data were therefore ordinal where a low score means that people

think A is worse than B and a high score means that people think B is worse than A. Both the test-retest reliability check and comparisons of the responses to the ‘generic’ and ‘contextual’ questions required within-subject tests of differences between responses. We therefore used Wilcoxon Signed Ranks tests to test whether the responses to the ‘generic’ and ‘contextual’ questions were significantly different from one another.

To test whether responses were significantly different when different ratios of numbers of deaths were presented required between-subject tests. We therefore used either Kruskal-Wallis (when three versions were compared) or Mann-Whitney tests (when two versions were compared) to test whether the responses to each question were significantly different between the versions of the questionnaire.

Thematic qualitative analysis was conducted on the written reasons that participants gave when they answered the ‘contextual’ questions. All responses were transcribed and coded by JC.

## **RESULTS**

### **1) Test-retest Reliability Check**

Table 2 summarises the results for the repeated ‘generic’ question that was asked twice in Part 1 of the questionnaire at Q6 and Q21. We report the percentages of respondents giving ratings from 1 (A is much worse than B) to 4 (B is much worse than A), the means and standard deviations of the ratings, and the Z statistic obtained from the Wilcoxon Signed Ranks test. Although the results show that the variance in responses and proportions rating B as ‘much worse’ than A appear to be slightly lower for Q21 than Q6 (i.e., standard deviations 0.86 vs.

0.99; 35.6% vs. 41.2%), the Wilcoxon Signed Ranks test showed that there was no significant difference in the distributions of responses between Q6 and Q21. The test-retest reliability of the ‘generic’ questions was therefore of an acceptable level.

TABLE 2 ABOUT HERE

## **2) Differences between responses to ‘generic’ and ‘contextual’ questions**

Table 3 summarises the results, aggregated across versions, for the ‘generic’ and ‘contextual’ questions for those pairings where the ratios of numbers were increased in the same direction in all versions of the questionnaire. We did not consider it meaningful to aggregate responses to questions where the numbers of deaths were ‘counterbalanced’ (but see Table 4 for the disaggregated results for these questions).

For each of the seven pairings from which we obtained responses from both ‘generic’ and ‘contextual’ questions we report the percentages of respondents giving ratings from 1 (A is much worse than B) to 4 (B is much worse than A), the means and standard deviations of the ratings, and the Z statistic obtained from the Wilcoxon Signed Ranks tests.

TABLE 3 ABOUT HERE

No significant differences in the distributions of responses were found for Accidents at work vs. Car Drivers – where the majority regarded deaths from Accidents at Work as worse than Car Driver deaths in both the ‘generic’ and ‘contextual’ versions (77.2% vs. 80.5%). Significant differences were, however, found for the remaining 3 pairings for which these comparisons could be made.

Car Drivers vs. Rail Passengers: Although similar percentages of respondents regarded rail passenger deaths as worse than car driver deaths in both the ‘generic’ and ‘contextual’ versions (80.0% and 82.4%), the ‘contextual’ version produced a lower proportion of ‘slightly worse than’ (28.8% vs. 44.4%) and higher proportion of ‘much worse than’ responses (53.6% vs. 35.6%).

Work-related Cancer vs. Car Drivers: Just over 20% more respondents regarded work-related cancer deaths as worse than car drivers deaths in the ‘contextual’ version than the ‘generic’ version (90.3% vs. 68.3%).

Smoking Cancer vs. Asbestos Cancer: About 10% more respondents regarded asbestos cancer deaths as worse than smoking cancer deaths in the ‘contextual’ version than the ‘generic’ version (73.4% vs. 69.2%) and the proportion of ‘much worse than’ responses were also higher (51.7% vs. 28.4%).

As indicated above, for three of the pairings: Pedestrians vs. Breast Cancer, Car Drivers vs. Pedestrians, CO Poisoning vs. Accidents at Work, the number dying by each cause were ‘counterbalanced’ across versions of the questionnaires (i.e. appeared both as 1.6:1 and 1:1.6). As it is not meaningful to aggregate responses to these questions, Table 4 therefore reports responses for these pairings at an individual question level.

Table 4 shows that in 2 of the 3 pairings considered there are significant differences between the ‘generic’ and ‘contextual’ responses. Only in the case of Pedestrians vs. Breast Cancer are the differences between the ‘generic’ and ‘contextual’ responses equivocal, with agreement between ‘contextual’ and ‘generic’ responses in one of the two pairings and disagreement in the other.

These findings lead us to the conclusion that people's responses to 'generic' questions in which the causes of death are represented by attributes only are not completely predictive of their responses to 'contextual' questions which provide additional information about the specific causes of death.

### **3) Sensitivity to different ratios of deaths in the 'generic' and 'contextual' questions**

For these tests the ratios of deaths in the questions were increased between the versions of the questionnaire depending on which pairing of scenarios was used (see Figure 2).

- (i) 1:1 vs. 1.5:1 vs. 2.5:1 – Car Drivers vs. Rail Accidents.
- (ii) 1.5:1 vs. 2:1 - Smoking Cancer vs. Asbestos Cancer; or 1:1.5 vs. 1:2 – Accidents at Work vs. Car Drivers and Work-related Cancer vs. Car Drivers.

Table 5 summarises the results for these comparisons. We report the percentages of respondents giving ratings from 1 (A is much worse than B) to 4 (B is much worse than A), the means and standard deviations of the ratings, and the chi-square ( $\chi^2$ ) or Z statistics obtained from the Kruskal-Wallis or Mann-Whitney U tests.

TABLE 5 ABOUT HERE

The results show that although participants' responses were equally sensitive or insensitive to the different ratios in both the 'generic' and 'contextual' questions in two of the four pairings (Car Drivers vs. Rail Passengers, Accidents at Work vs. Car Drivers), sensitivity to the ratios was affected by providing information about the specific causes of death in the other two pairings. However, the way in which providing this information affected sensitivity was not the same. For one pairings the sensitivity seems to have been greater in the 'contextual'

questions than the ‘generic’ questions (Work-related Cancer vs. Car Drivers), with the opposite result for the other pairing (Smoking Cancer vs. Asbestos Cancer).

Overall, there was therefore no systematic tendency for the responses to the ‘generic’ questions to be any more or less valid than those to the ‘contextual’ questions. So we cannot draw any definitive conclusion about which set of responses is by default the most appropriate.

However, these results also suggest that any conclusions we might draw about the relative importance that people place on the different attributes from the responses they gave to the ‘generic’ questions might be quite different from the conclusions we might draw from their responses to the ‘contextual’ questions. So if we take these results along with those reported in the previous section it appears that people’s responses to the ‘generic’ questions do not serve as reliable proxies for their responses to ‘contextual’ questions. The reasons for this mis-match were explored by analyzing the written reasons that participants gave when they answered the ‘contextual’ questions.

#### **4) Results of the thematic qualitative data analysis**

Thematic analysis was conducted on the written reasons that participants gave when they answered the ‘contextual’ questions. We were particularly interested in gaining insight into the reasons why their responses to these questions were different to the ‘generic’ equivalents. Hence, the analysis was only conducted on the five pairings shown in Table 3 and Table 4 which produced significant differences between the ‘generic’ and ‘contextual’ questions – focusing our attention on the reasons given by those participants who gave very different

answers to both questions (i.e., rated A as worse than B in the ‘generic’ version and rated B as worse than A in the ‘contextual’ version – or vice versa).

#### Car Drivers vs. Rail Passengers

28 of the respondents who switched to rating B (rail passenger deaths) as worse than A (car driver deaths) in the ‘contextual’ version gave written comments. The majority of the reasons given were linked to the ‘blame’ and ‘quality of life’ attributes that were provided in the ‘generic’ questions – i.e., that the individuals were not to blame or were helpless (14 cases – *“Not the individual’s fault” [1005\_v1]; “The individual is powerless to influence events” [1048\_v3]*), the greater suffering of the rail accident victims (6 cases – *“Due to the suffering that they had more than to who was to blame” [2063\_v3]*), or the violent nature of the death in a rail accident (2 cases – *“In rail accidents parts of bodies are found everywhere along the railway lines” [1023\_v2]*).

However, a notable minority of reasons provided seemed to be beyond the scope of the attributes presented in the ‘generic’ questions – issues that expressed the view that deaths should just not be happening on the railways – i.e., the railways should be safe (10 cases *“Whilst all travel involves risk, when traveling by train danger is not expected” [1057\_v3]*), and rail accidents should be avoidable (2 cases – *“Avoidable by better maintenance of track and carriage” [2044\_v3]*).

#### Smoking Cancer vs. Asbestos Cancer

46 of the respondents who switched to rating B (asbestos cancer) as worse than A (smoking cancer) in the ‘contextual’ version gave written comments. All of the themes that emerged were linked to the ‘blame’ attribute – i.e., that smokers only have themselves to blame

whereas the victims of asbestos related cancer were not at fault (29 cases “*Smoking is a choice so if they are dying from smoking related cancer then its only themselves to blame*” [1003\_v1]; “*It was not their fault*”[2095\_v3]), are unaware of the damage that asbestos might cause (17 cases “*People were not warned of asbestos damage it could cause*” [2065\_v3]), or that business and government should take responsibility for people’s safety regarding asbestos (9 cases “*It is the duty of business/ government to lay down guidelines concerning asbestos*” [1048\_v3]).

### Car Drivers vs. Pedestrians

38 of the respondents who switched to rating B (pedestrian deaths) as worse than A (car driver deaths) in the ‘contextual’ version gave written comments. Most of the themes that emerged were linked to the ‘blame’ (23 cases “*Due to the fact that pedestrians are not to blame*” [1073\_v1]), ‘age’ (2 cases “*I think this is a vulnerable age group in this type of accident and should be given more consideration*” [2104\_v3]), and ‘number of deaths’ attributes (6 cases “*I feel B is slightly worse because of the number of deaths*” [1029\_v3]).

However, a notable minority of reasons provided seemed to be beyond the scope of the attributes presented in the ‘generic’ questions – issues that expressed the view that pedestrians have the right to feel safe walking the streets (4 cases “*It should be safe to walk the streets*” [2024\_v1]) or mentioned the impact of the deaths on other people (2 cases “*Although all car accident related deaths are tragic to those concerned, it must be especially difficult for those grieving when there is someone else who can be held accountable*” [2060\_v1]).

### Work-related Cancer vs. Car Drivers



28 of the respondents who switched to rating A (work-related cancer deaths) as worse than B (car driver deaths) in the ‘contextual’ version gave written comments. The majority of written comments were linked to the ‘blame’ and ‘quality of life’ attributes – i.e., the work-related cancer victims had no choice over the situation and their deaths were caused by others (14 cases *“The individuals had no control”* [2027\_v1]; *“Awful, why so many deaths, again due to business or government. In scenario B, the car drivers are to blame for the deaths”* [2113\_v2]), or that the work-related cancer victims had greater or longer suffering (10 cases *“I think A is much worse than B because suffering one-two years before death is more painful than to die instantaneously”* [2045\_v2]).

More generally however, a number of respondents expressed the view that workers should not be put at risk in the work-place (10 cases *“You should not be put at risk in a work place and should be protected by law”* [2058\_v1]), or that these deaths could have been prevented (3 cases *“Cancer could have been prevented if they hadn’t been exposed to the chemicals at work due to the business”* [1087\_v2]).

#### CO Poisoning vs. Accidents at Work

33 of the participants who switched to rating A (CO poisoning deaths) as worse than B (deaths in accidents at work) in the ‘contextual’ version gave written comments. The reasons given for this pairing were only indirectly linked to the ‘blame’ and ‘quality of life’ attributes. They tended to focus on issues of avoidability and neglect rather than who was to blame (21 cases *“Easily avoidable if government funded installing carbon monoxide detectors in homes. These deaths are totally unnecessary”* [1128\_v3]; *“Gas appliances should be checked, this is neglect”* [2038\_v3]), or the nature of the death from poisoning (2 cases *“Poisoning in any way or form is in my opinion one of the worst deaths imaginable”* [1067\_v2]).

In summary, the qualitative data highlight two main reasons why the ‘contextual’ questions produce different responses to the ‘generic’ questions.

The first reason is that the five attributes used to describe the causes of death in the ‘generic’ questions may not fully capture the differences between the deaths that participants want to take into account when giving their ratings. For example, in the comparison between car drivers and rail passengers some of the reasons given picked up on the fact that rail deaths were regarded as worse because participants expect the railways to be a safe place or that rail accidents are avoidable. The expectation of a safe environment also came into play with the pedestrian deaths, and the avoidability issue also came into play with the CO poisoning deaths and work-related cancer deaths. The other main issue that may not have been captured by the quality of life attribute was some participants’ perceptions of the particularly nasty nature of the rail accident and CO poisoning deaths.

The second reason is that the interpretation of the levels used to describe the attributes may cover quite a wide range of different situations. This is most clearly illustrated by the ‘individuals themselves’ and ‘business or government’ levels of the ‘blame’ attribute both of which differentiated between the Car Accident vs. Rail Passenger and Smoking Cancer vs. Asbestos Cancer pairings. However, it is notable that whereas in the Car Accident vs. Rail Passenger pairing more of the reasons for rating the rail passenger deaths as worse than the car accident deaths were related to the fact that rail accident victims were not to blame, in the Smoking Cancer vs. Asbestos Cancer pairing more of the reasons for rating the asbestos cancer deaths as worse than the smoking cancer death were related to the fact that smokers were to blame. This suggests that the level ‘business or government’ was regarded as more

deserving of a 'worse than' rating when it was used to describe rail passenger deaths than when it was used to describe asbestos cancer deaths, and similarly the level 'individuals themselves' was regarded as less deserving of a 'worse than' rating when it was used to describe smoking cancer deaths than car driver deaths. In other words although participants will have accepted that 'business or government' was most to blame in the cases of asbestos cancer deaths and rail passenger deaths, the actual degrees of business or government responsibility might be perceived quite differently by participants for these two types of deaths. Similarly participants might perceive the actual degrees of individual responsibility associated with car driver deaths and smoking cancer deaths quite differently.

## **Discussion**

Our study sought to establish the extent to which 'generic' descriptions where causes of death are described merely in terms of their standing on a number of key attributes (age, severity and duration of pain and suffering, blameworthiness) are predictive of the response to 'contextual' descriptions where the causes of death are identified. We find evidence of differences in responses between the 'generic' and 'contextual' questions, but no clear pattern as to the direction of these differences.

The qualitative data suggest that two factors led to differing perceptions between the 'generic' and 'contextual' questions. Firstly, some influential variable(s) were omitted from the 'generic' questions (for example, the 'violent' nature or 'avoidability' of deaths on the railways). It is possible that the use of a larger number of attributes may have overcome this problem. The authors, however, acknowledge the tension between the need to describe the totality of influencing factors and the need to arrive at a manageable set of attributes.<sup>(12)</sup>

Secondly, when the causes of death were provided certain categories of included variables were interpreted somewhat differently (for example, greater importance seemingly being attached to the ‘blame’ dimension in some cases). Whilst the problem of omitted variables can be solved at least conceptually, there is no obvious means of overcoming the finding that the interpretation of attributes is context dependent. Thus, our study highlights the difficulty of trying to use a ‘generic’ set of attributes to anticipate people’s responses towards a set of attributes where the specific cause of death is identified. But this, in itself, does not allow us to say that one set of values are superior to the other.

As set out above, ‘sensitivity to theoretically relevant factors’ – such as the numbers dying – is one criterion against which the validity of responses may be measured (Loomes<sup>(10)</sup>). Whilst we found differences in the sensitivity to the numbers of deaths between the ‘contextual’ and ‘generic’ questions, there was no clear pattern to this finding. Hence, it is difficult to point to any empirical support for the superiority of one set of responses over another.

To explore this issue further, let us consider more carefully the factors that underpin decision making. Loewenstein and O’Donoghue<sup>(13)</sup> make a distinction between ‘affective’ system and ‘deliberative’ system for decision making – where affective are based on emotive impressionistic reactions, and deliberative involve more systematic weighing up of consequences. For example, Slovic et al.<sup>(14,15)</sup> argue that, in responding to questions involving contexts that carry strong negative affect meanings, respondents are less sensitive to probability information than contexts that carry less affect. For example, with hazards like nuclear power and exposure to small amount of toxic chemicals, the negative consequences of these risks may make respondents more concerned and sensitive to the possibility of these

risks, and less sensitive to information about the actual probability of these risks. It could be argued that, in such cases, the affective responses – based on emotive impressionistic considerations – are ‘distorting’ responses and ought to be ‘factored out’.

Likewise, it is often considered in the health state valuation literature that values for health states ought not to take into consideration ‘emotive’ issues, such as those surrounding cancer. Rather, health states are generally described in terms of ‘generic’ health state classification systems such as the EQ 5D, which involves 3 levels on 5 dimensions (mobility, self-care, usual activities’, pain, anxiety & depression).<sup>(16)</sup> Leaving aside issues concerning the crude nature of the classification systems (and other criticisms of the EQ 5D and other ‘generic’ measures), such systems aim to achieve comparability across the range of health care interventions in order to aid resource allocation decisions. It has long been argued (see, for example Drummond et al.<sup>(17)</sup>) that such decisions ought not to be based on the aggregation of the ‘disease specific’ preferences of disparate patient groups as the values are incompatible.

This brings us to the fundamental point that the superiority of ‘generic’ or ‘contextual’ valuations may ultimately depend upon the purpose to which the responses are put. If there was a desire to use people’s responses about which deaths are particularly bad to inform regulatory policy there is certainly something appealing about the health economics argument for ‘generic’ measures in order to aid comparability in resource allocation decisions. It may also seem desirable to reject ‘affective’ in favour of ‘deliberative’ responses. On the other hand, the results of our study indicate that the public may well reject the policy implications that arise from their own responses to ‘generic’ questions of the type posed here. Likewise, the public may well have a genuine desire to save cancer sufferers for reasons that can never be captured in any ‘generic’ system, however fully specified.

In his review of the literature of value elicitation in the fields of health, safety and the environment Loomes<sup>(10)</sup> suggested that policy makers should generally try to elicit a broad list of principals on which to allocate resources from respondents. The implications of these rules should be also shown to respondents, to ensure that they agree with the implications and that they cannot find any way to improve well-being. In this context, the feedback mechanism would certainly involve a discussion of the context and example of the types of decisions the general rules would dictate. Were respondents to overwhelmingly reject the implications of general rules once the implications of the rules are made clear, the rules would need to be revised. This approach suggests the need for an iterative procedure, similar to a citizen's jury – where rules are revised and adjusted by a panel of lay people, that are consulted and reconsulted about appropriate decision rules. Whether the 'gap' between 'contextual' and 'generic' rules can be narrowed by such an iterative process has yet to be explored.

## References

1. Sunstein, C.R. (1997). Bad deaths. *Journal of Risk and Uncertainty*, 14(3), 259-282.
2. Jones-Lee, M.W., Hammerton, M., Philips, P.R., (1985). The value of safety- results from a national sample survey. *Economic Journal*, 95(377), 49-72.
3. Savage, I. (1993). An empirical-investigation into the effect of psychological perceptions on the willingness-to-pay to reduce risk. *Journal of risk and uncertainty*, 6(1), 75-90.
4. Subramanian, U., & Cropper, M. (2000). Public choices between life saving programs: The tradeoff between qualitative factors and lives saved. *Journal of Risk and Uncertainty*, 21(1), 117-149.
5. Slovic, P. (1992). Perceptions of risk: Reflections on the psychometric paradigm. In S. Krimsky, and D. Golding. (Eds). *Social Theories of Risk*. Westport, CT: Praeger, pp. 117-152.
6. Sackett, D.L., Torrance, G.W.(1978). Utility of different health states as perceived by general public. *Journal of chronic diseases*, 31(11), 697-704.
7. Gerard, K., Dobson, M., and Hall, K. (1993). Framing and labeling effects in health descriptions: Quality adjusted life years for treatment of breast cancer. *Journal of Clinical Epidemiology*, 46, 77-84.
8. Rabin, R., Rosser. R.M., & Butler, C. (1993). Impact of diagnosis on utilities assigned to states of illness. *Journal of the Royal Society of Medicine*, 86, 444-448.
9. Robinson, S., Bryan, S., (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, University of Oxford.
10. Loomes, G. (2006). (How) Can we value health, safety and the environment? *Journal of economic psychology*, 27(6), 713-736.
11. Robinson A, Chilton S, Covey J, Jones-Lee M, Loomes, G, Spencer A. (2007). Valuation of health and safety benefits Dread risks. Health and Safety Executive.
12. Coast J, Horrocks S. (2007). Developing attributes and levels for discrete choice experiments using qualitative methods. *Journal of Health Services Research & Policy*, 12, 25-30.
13. Loewenstein, G., & O'Donoghue, T.(2004). Animal spirits. Affective and deliberative processes in economic behavior, Mimeo.
14. Slovic, P., Finucane, M.L., Peters, E., MacGregor, D.G. (2007). The affect heuristic *European journal of operational research*, 177(3), 1333-1352.
15. Slovic, P., Finucane, M.L., Peters, E., MacGregor. D.G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, 24(2), 311-322.
16. Dolan, P., Gudex, C., Kind, P. & Williams, A. (1996). The time trade-off method: results from a general population study. *Health Economics*, 5, 141-154.
17. Drummond, M.F., O'Brien, B.J., Stoddart, G.L., Torrance, G. (1997). *Methods for the economic evaluation of health care programmes*. 2nd ed. Oxford University Press, Oxford.

## **Acknowledgements**

The research reported in this paper was carried out under a project funded by the Health and Safety Executive. However, the opinions expressed in this paper are solely the responsibility of the authors and do not necessarily reflect the views of the project sponsor. The authors would like to thank Aki Tsuchiya for discussing an earlier version of this paper at a Health Economics Study Group Meeting held in Newcastle in July 2005.



**Table 1: Demographic characteristics of the sample**

	Frequency(Percent)
Gender	
Male	154(49.2)
Female	159(50.8)
Age	
17-34	121(38.7)
35-54	97(31.0)
55-90	94(30.0)
Missing	1(0.3)
long-term illness, health problem or disability that limits daily activity or work	
No	253(80.8)
Yes	54(17.3)
Missing	6(1.9)
Members of your household with long-term illness, health problem or disability that limits daily activity or work	
No	279(89.1)
Yes	28(8.9)
Missing	6(1.9)

**Table 2: Test-retest reliability check on repeated 'generic' question**

A vs. B	% Response				Mean	SD	Z <sup>a</sup> (p)
	1	2	3	4			
Repeated 'generic' Question (N=306)							
(Q6)	10.8	13.1	35.0	41.2	3.07	0.99	Z=0.61
(Q21)	6.5	13.2	44.4	35.6	3.09	0.86	(p=.54)

Notes:

<sup>a</sup>Wilcoxon Signed Ranks Test

**Table 3: Comparison of responses to ‘generic’ and ‘contextual’ questions**

A vs. B	% Response				Mean	SD	Z <sup>a</sup> (p)
	1	2	3	4			
Car Drivers vs. Rail Passengers (N=306)							
Generic (Q21)	6.5	13.2	44.4	35.6	3.09	0.86	Z=3.48
Contextual	7.5	10.1	28.8	53.6	3.28	0.93	(p<.001)
Work-related Cancer vs. Car Drivers (N=205)							
Generic	31.7	36.6	20.5	11.2	2.11	0.98	Z=7.81
Contextual	61.5	28.8	8.8	1.0	2.49	0.70	(p<.001)
Smoking Cancer vs. Asbestos Cancer (N=201)							
Generic	20.4	20.4	30.8	28.4	2.67	1.10	Z=4.31
Contextual	14.4	12.4	21.4	51.7	3.10	1.10	(p<.001)
Accidents at Work vs. Car Drivers (N=206)							
Generic	37.4	39.8	16.5	6.3	1.92	0.89	Z=0.81
Contextual	42.2	38.3	10.7	8.7	1.86	0.93	(p=0.42)

Notes:

<sup>a</sup>Wilcoxon Signed Ranks Test

**Table 4: Comparison of responses to ‘generic’ and ‘contextual’ questions disaggregated for three pairings.**

		% Response							
A vs. B		1	2	3	4	Mean	SD	$\chi^2$ <sup>a</sup> or Z <sup>b</sup> (p)	
Pedestrians vs. Breast Cancer									
Generic	1.6:1	16.0	19.8	39.6	24.5	2.73	1.01	Z=1.39	
	1:1.6	11.0	19.0	37.0	33.0	2.92	0.98	(p=.16)	
‘contextual’	1.6:1	17.0	30.2	33.0	19.8	2.56	1.00	Z=3.27	
	1:1.6	8.0	17.0	42.0	33.0	3.00	0.91	(p<.001)	
Car Drivers vs. Pedestrians									
Generic	1.6:1	19.0	30.0	36.0	15.0	2.47	0.97	Z=4.19	
	1:1.6	8.1	17.2	36.4	38.4	3.05	0.94	(p<.001)	
‘contextual’	1.6:1	11.0	15.0	45.0	29.0	2.92	0.94	Z=5.00	
	1:1.6	1.0	4.0	35.4	59.6	3.53	0.63	(p<.001)	
CO Poisoning vs. Accidents at Work									
Generic	1.6:1	12.3	29.2	38.7	19.8	2.66	0.93	Z=8.47	
	1:1.6	0.0	1.0	25.5	73.5	3.72	0.47	(p<.001)	
‘contextual’	1.6:1	24.5	30.2	38.7	6.6	2.27	0.91	Z=6.11	
	1:1.6	10.2	10.2	37.8	41.8	3.11	0.96	(p<.001)	

Notes:

<sup>a</sup>Kruskall-Wallis Test, <sup>b</sup>Mann-Whitney U Test

**Table 5: Sensitivity to ratios of deaths in ‘generic’ and ‘contextual’ questions**

A vs. B		% Response				Mean	SD	$\chi^2$ <sup>a</sup> or Z <sup>b</sup> (p)
		1	2	3	4			
Car Drivers vs. Rail Passengers								
Generic (Q21)	1:1	5.0	8.9	37.6	48.5	3.30	0.83	$\chi^2=11.69$ (p=.003)
	1.5:1	3.7	13.1	54.2	29.0	3.08	0.75	
	2:1	11.2	18.4	40.8	29.6	2.89	0.96	
‘contextual’	1:1	7.9	5.0	21.8	65.3	3.45	0.91	$\chi^2=8.19$ (p=.017)
	1.5:1	4.7	9.3	37.4	48.6	3.30	0.83	
	2:1	10.2	16.3	26.5	46.9	3.10	1.02	
Work-related Cancer vs. Car Drivers								
Generic	1:1.5	38.4	34.3	17.2	10.1	1.99	0.98	Z=1.87 (p=.062)
	1:2	25.5	38.7	23.6	12.3	2.23	0.97	
‘contextual’	1:1.5	69.7	23.2	6.1	1.0	1.38	0.65	Z=2.34 (p=.019)
	1:2	53.8	34.0	11.3	0.9	1.59	0.73	
Smoking Cancer vs. Asbestos Cancer								
Generic	1.5:1	14.1	15.2	33.3	37.4	2.94	1.05	Z=3.45 (p<.001)
	2:1	26.5	25.5	28.4	19.6	2.41	1.08	
‘contextual’	1.5:1	14.1	14.1	19.2	52.5	3.10	1.11	Z=0.05 (p=.96)
	2:1	14.7	10.8	23.5	51.0	3.11	1.10	
Accidents at Work vs. Car Drivers								
Generic	1.5:1	34.0	42.5	20.8	2.8	1.92	0.81	Z=0.62 (p=.54)
	2:1	41.0	37.0	12.0	10.0	1.91	0.96	
‘contextual’	1.5:1	44.3	38.7	11.3	5.7	1.78	0.86	Z=0.96 (p=.34)
	2:1	40.0	38.0	10.0	12.0	1.94	0.99	

Notes:

<sup>a</sup>Kruskall-Wallis Test, <sup>b</sup>Mann-Whitney U Test

**Figure 1. Example of ‘generic’ question**

**Which is worse?**

	<b>A</b>	<b>B</b>
<b>Number of people who die</b>	15 deaths	25 deaths
<b>Age-group</b>	Over 60 year olds	Over 60 year olds
<b>Quality of life in period leading up to death</b>	A bit worse than normal for last 1-2 years of their lives	A lot worse than normal for last 1-2 years of their lives
<b>Who is most to blame</b>	The individuals themselves	Business or Government

What do YOU think?  
(tick one)

A is much  
worse than B

A is slightly  
worse than B

B is slightly  
worse than A

B is much  
worse than A

**Figure 2: Causes of death used in the ‘contextual’ questions in each version of the questionnaire**

	Version 1	Version 2	Version 3
A: Car Drivers	10 deaths	15 deaths	25 deaths
B: Rail Passengers	10 deaths	10 deaths	10 deaths
A: Pedestrians	15 deaths	25 deaths	
B: Breast Cancer	25 deaths	15 deaths	
A: Work-related Cancer	10 deaths	25 deaths	
B: Car Drivers	15 deaths	50 deaths	
A: Car Drivers	25 deaths		15 deaths
B: Pedestrians	15 deaths		25 deaths
A: Smoking Cancer	50 deaths		15 deaths
B: Asbestos Cancer	25 deaths		10 deaths
A: Accidents at Work		10 deaths	25 deaths
B: Car Drivers		15 deaths	50 deaths
A: CO Poisoning		25 deaths	15 deaths
B: Accidents at Work		15 deaths	25 deaths

**This working paper has been produced by  
the Department of Economics at  
Queen Mary, University of London**

**Copyright © 2007 Anne Spencer, Judith Covey, Angela Robinson  
and Graham Loomes. All rights reserved**

**Department of Economics  
Queen Mary, University of London  
Mile End Road  
London E1 4NS  
Tel: +44 (0)20 7882 5096  
Fax: +44 (0)20 8983 3580  
Web: [www.econ.qmul.ac.uk/papers/wp.htm](http://www.econ.qmul.ac.uk/papers/wp.htm)**