Journal of Risk and Uncertainty, 17:5–25 (1998)

On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1-*Caveat Investigator*

JANE BEATTIE*
University of Sussex

JUDITH COVEY
University of York

PAUL DOLAN
University of Sheffield

LORRAINE HOPKINS
University of Wales, Bangor

MICHAEL JONES-LEE
University of Newcastle upon Tyne

GRAHAM LOOMES
University of Newcastle upon Tyne

NICK PIDGEON
University of Wales, Bangor

ANGELA ROBINSON
University of Newcastle upon Tyne

ANNE SPENCER
University of York

Abstract

This article reports the results of two studies aimed at testing and refining a procedure for estimating willingness-to-pay based monetary values of safety using the contingent valuation method. In spite of the fact that respondents were given the opportunity to discuss various safety issues and key concepts in focus group meetings held in advance of individual interviews, and were also given ample opportunity to revise their responses in the light of the overall pattern of these responses, the results show clear evidence of extensive and persistent insensitivity to the scale and scope of the safety improvements that were specified in the contingent valuation questions, as well as vulnerability to framing effects. This clearly casts serious doubt on the reliability and validity of willingness-to-pay based monetary values of safety estimated using conventional contingent valuation procedures.

Key words: safety, contingent valuation, embedding, scope, sequencing.

JEL Classification: J17

Under the so-called "willingness-to-pay" (WTP) approach to the valuation of safety, the monetary value of a particular safety improvement is defined as the (possibly weighted) aggregate of the amounts that those individuals affected by the improvement would be willing to pay for it. As such, the WTP approach is firmly rooted in the fundamental prescriptive premise of conventional social cost-benefit analysis which requires that public sector allocative and regulatory decisions should, so far as possible, reflect the preferences of those who will be affected by the decision concerned. It is therefore not surprising that this approach to the valuation of safety is enjoying increasing currency both in the UK and abroad and is, for example, explicitly recommended by the UK Treasury and the US Department of Transportation.[1]

To date, "direct" empirical estimates of WTP-based values of safety have been obtained by methods that can be broadly classified into one of two types, namely "revealed preference" and "contingent valuation".[2] Essentially, under the revealed preference approach, WTP-based values are elicited from data concerning *actual* choices involving explicit or implicit trade-offs of wealth for risk. By contrast, the contingent valuation (CV) approach involves asking members of a representative sample of the population at risk more or less directly about their willingness to pay for (typically small) *hypothetical* improvements in their own (and possibly other people's) safety.

In addition to direct empirical estimation procedures, "indirect" methods have also been used to estimate the monetary value of preventing one severity or type of harm *relative* to the value of preventing another. Thus, for example, the UK Department of the Environment, Transport and the Regions' (DETR) current monetary values for the prevention of different severities of non-fatal road injury were obtained by applying such relative valuations to an absolute monetary "peg" in the form of the DETR WTP-based monetary value for the prevention of a road fatality.[3] Prominent amongst the relative valuation approaches are those involving marginal "risk-risk" (RR) questions, "standard gamble" (SG) questions and so-called "matching" (or "equivalence") questions — for a more detailed account of the nature of RR, SG and matching questions, see for example Viscusi *et al* (1991) Jones-Lee *et al* (1995) and Jones-Lee and Loomes (1995). To the extent that all three of these relative valuation approaches ask a representative sample of the population at risk questions involving hypothetical choices or trade-offs, they clearly have very much more in common with the CV approach than with revealed preference.

Were it to be the case that for any particular severity of harm, a single "universally transferable" WTP-based value of safety was applicable to allocative and regulatory decision making in all contexts (such as roads, rail, the workplace, power generation and so on), then it would in principle be possible to obtain an estimate of the relevant value using *either* of the two direct empirical estimation procedures, provided that the appropriate data were available. Thus, a WTP-based value for the prevention of a fatality obtained from, say, workplace data using the revealed preference approach could also be employed in the context of road safety, or in the appraisal of risks from nuclear power generation.

Unfortunately, however, there is growing evidence that WTP-based values of safety are *not* universally transferable and that people's *ex ante* willingness to pay to reduce risk will instead tend to vary with their perceptions of and attitudes towards the characteristics of different hazards, such as the extent to which the hazard concerned is seen to be volun-

tarily assumed, under potential victims' own control, their own responsibility, well-understood, and so on.

Thus for example, Mendeloff and Kaplan (1990) found that median willingness to pay to prevent a given number of deaths varied by up to a factor of more than three across various different contexts, such as workplace exposure to carcinogenic chemicals, bicycle and automobile accidents and fatal crib-slat accidents to young children. Similarly, M^{c}Daniels *et al* (1992) found that for more familiar and well-defined hazards, such as automobile and aviation accidents, individual willingness to pay to reduce risk was most substantially influenced by respondents' perceived personal exposure to the hazard concerned. By contrast, for less familiar and more poorly understood hazards, such as nuclear power and electromagnetic fields, the most important influences were levels of "dread" and the perceived severity of adverse consequences. In turn, Savage (1993) found substantial differences in mean willingness to pay to reduce the risks from road and aviation accidents, domestic fires and stomach cancer, with willingness to pay being significantly affected by various psychological factors including perceptions of the "dread" and "unknown" attributes of the hazard concerned. Finally, on the basis of a study carried out in the London area, Jones-Lee and Loomes (1995) found that, on average, respondents valued the prevention of a fatality on London Underground at about one and a half times the value that they placed on the prevention of a road fatality.[4]

In view of this likely variability across different contexts, a consortium of UK government departments[5] commissioned a program of research—to be undertaken by some of the authors of this article—aimed at estimating a "tariff" of WTP-based values of safety for a number of different contexts, including the roads and other public transport modes; the workplace; domestic fires and nuclear power.[6] For various reasons, including lack of the requisite data, it was decided that the revealed preference approach would *not* be feasible in most of these contexts and that a combination of the CV and relative valuation approaches should therefore be employed.

More specifically, the aim was to use direct CV questions to estimate absolute monetary values of safety in those contexts (such as the roads) in which it was felt that base risks were "large" enough to render CV questions concerning reductions in risk comprehensible to respondents. By contrast, in other contexts (such as rail or nuclear power generation) it would appear that base risks are so low that CV questions concerning reductions in those risks would involve probabilities that most members of the public would find great difficulty in conceptualising. In addition, the calculation by which CV responses are processed to produce monetary values of safety is such that even small errors in the responses themselves would be "blown up" into unacceptably wide error bands on the monetary values concerned.[7] For these low base risk contexts it was therefore decided to use relative valuation questions in order to estimate values of safety in those contexts relative to the corresponding values in contexts in which the CV approach could be applied directly.

A further reason for using the CV and relative valuation approaches—rather than revealed preference—was that in principle these approaches allow an in-depth exploration of the thought processes by which people arrive at decisions concerning the trade-offs involved in the questions. Such in-depth exploration was felt to be particularly important

in view of the growing evidence of apparent anomalies and inconsistencies in responses to CV and other preference-elicitation questions in the safety and environmental fields. Amongst the more worrying of these anomalies and inconsistencies are so-called "embedding", "scope" and "sequencing" effects. Essentially, embedding and scope effects refer to the tendency of many CV respondents to report much the same willingness to pay for a comprehensive bundle of safety or environmental "goods" as for a proper subset of that bundle. In turn, sequencing effects reflect a tendency for the order in which a sequence of CV questions are put to respondents to have a significant impact on the values that are implied by the responses to such questions.[8]

With these considerations in mind it was decided to conduct a fairly extensive and detailed program of work developing and refining the survey instruments prior to commencement of the main fieldwork. The purpose of this paper is to report the CV results of the first two phases of this research program which took place between October 1995 and July 1996 and involved a total of 135 respondents.

1. The first phase study

Because we wished not only to explore the thought processes by which people arrive at responses to safety-related CV questions, but also to give respondents every opportunity to reflect upon and clarify their thinking about difficult and often unfamiliar money/risk trade-offs, we elected to adopt a three-stage design in each of the first two phases of the research program. The three stages were as follows:

(i) Respondents were first recruited to focus groups of 5 or 6 participants. In these focus groups (moderated by members of the research team) various safety issues were discussed and participants were introduced to the stimuli and concepts that were to be used in the individual interviews which were to follow in the second stage.
(ii) One-to-one interviews were then conducted (typically within a few days of the initial focus group meetings) in which respondents completed a specially designed questionnaire containing a variety of direct CV and indirect relative valuation questions. All interviews were carried out by four members of the research team who had been involved in the design and pre-testing of the questionnaires.
(iii) Feedback meetings were then arranged (again, typically involving 5 or 6 participants) in order to show respondents the patterns of results they had collectively generated and to invite their reactions and comments. These meetings also provided respondents with the opportunity for further reflection and comment upon the nature of the thought processes underlying their CV responses.

The adoption of this intensive study design necessarily meant that the numbers of respondents in each of the first two phases of the research program were relatively small. Hence, the studies were designed in order to allow as many within-subject (as opposed to

between-sample) tests of consistency as possible. Throughout, respondents were encouraged to articulate the thought processes underlying their responses and all focus group discussions, individual interviews and feedback meetings were tape-recorded.

1.1 Design of the first phase study

The CV questions in the first phase study asked respondents about their willingness to pay for small reductions in their own risk of death or injury in a road traffic accident. These questions were constructed in such a way as to test whether CV responses, when elicited within the type of study design outlined above, were (a) sensitive to the magnitude and scope of the risk reduction and (b) independent of the sequence in which the questions were asked.

A total of 83 respondents from the York and Newcastle areas took part in the first phase study. This was essentially a convenience sample recruited through local organisations and from the relatives of pupils at local schools, but was nonetheless broadly representative in terms of age, gender and household income.

1.2 Risk stimuli used

Brief descriptions of four different severities of injury, which effectively span the spectrum of those that may occur as a result of a road traffic accident, were printed on small cards. While these injury descriptions are reproduced in full in the Appendix, sylised descriptions, along with associated annual probabilities of occurrence for the typical car driver or passenger in the UK, were as follows:

Code Letter	Injury Severity	Annual Risk
F	Fatal	6 in 100,000
P	Serious (Permanent)	20 in 100,000
T	Serious (Temporary)	50 in 100,000
M	Minor	500 in 100,000

In order that respondents could be given a graphical representation of the magnitude of these risks, a "risk showcard" was produced for each of the 4 injury severities. Each card showed a grid containing 100,000 white dots and resondents were told that each represented a "typical" driver or passenger. In the top left hand corner of this grid a number, x, of dots had been coloured in to represent the x in 100,000 drivers or passengers being killed/injured each year. These risk showcards, along with their associated injury descriptions, were shown to respondents towards the end of the initial focus group meetings. Thus, respondents had the opportunity to discuss and clarify the way in which the risk information was presented well in advance of the individual interviews.[9] This was done in light of the finding that many people appear to have difficulty in processing information

concerning small probabilities (Kahneman *et al* (1982)). However, respondents were not, at this stage, given any information relating to risk *reductions* nor were they introduced to the particular CV questions they were to be asked during the subsequent individual interviews.

1.3 Prioritisation exercise

Prior to answering the CV questions in the individual interviews, respondents were given a set of five cards each describing a safety improvement in the form of a reduction in the annual risk of one of the injury types described above. In order to test their sensitivity to the magnitude of the risk reduction, respondents were presented with two different reductions in their annual risk of death, namely 1 in 100,000 and 3 in 100,000, referred to in what follows as [F1] and [F3] respectively.

The five risk reductions and their associated code letters were as follows:

[F1]	1 in 100,000 reduction in annual risk of death
[F3]	3 in 100,000 reduction in annual risk of death
[P]	10 in 100,000 reduction in annual risk of injury P
[T]	25 in 100,000 reduction in annual risk of injury T
[M]	250 in 100,000 reduction in annual risk of injury M

Respondents were asked to prioritise these safety improvements from 1 to 5 such that the one they considered to be their top priority was number 1, etc. This was primarily intended as a "warm up" exercise, enabling respondents to familiarise themselves with the risk reduction information prior to answering the CV questions. The median ranking from top to bottom priority was: [F3], [P], [F1], [T], [M].

1.4 CV questions

After prioritising the risk reductions, respondents were presented with a series of 5 CV questions in which they were asked how much they would be willing to pay to have one or more of the 5 risk reductions implemented. The scenario they were asked to imagine was that they could have a safety feature fitted to their car which would reduce their own risk of death or injury but would leave the risk to other occupants of the car unchanged. Respondents were also told that if it was to remain effective, the safety feature would have to be renewed every twelve months. The first phase CV questions therefore focused on individual willingness to pay for a "self only", one-year risk reduction.

In order to test whether CV responses were sensitive to the scope of the safety improvement we decided to include questions relating both to a particular risk reduction on its own and to more inclusive "bundles" of reductions. Further, in order to test for sequencing effects, two variants of the questionnaire were produced, referred to henceforth

as "Top Down" (TD) and "Bottom Up" (BU). The TD version began with a question in which all risks were reduced simultaneously, followed by a sequence of questions in which successively smaller subsets of these risks were reduced, while the BU sequence started with a question about a reduction in the risk of death on its own and ended with a question in which all risks were reduced simultaneously. Thus, the respective sequences of CV questions put to the TD and BU groups were as follows:

Group	Q1	Q2	Q3	Q4	Q5
TD	[F3+P+T+M]	[F3+P+T]	[F3+P]	[F3]	[F1]
BU	[F1]	[F3]	[F3+P]	[F3+P+T]	[F3+P+T+M]

All CV questions were printed on a left hand page of the questionnaire with, on the facing right hand page, a response scale rising in equal increments of £1 from £0 to £50 and then in equal increments of £10 from £50 to £550, together with an "over £550" response.

Respondents were then asked to put a tick next to each amount they were sure they *would* be prepared to pay for the safety feature, to put a cross next to each amount they were sure they would *not* be prepared to pay and, finally, to put an asterisk next to the amount at which they would find it most difficult to decide whether or not to pay that much for the safety feature[10]. After completing their set of five CV questions, respondents were presented with their asterisked responses and asked to reflect upon their series of answers. Respondents were then told that they could alter any, or all, of their initial responses if they so wished.

1.5 CV results from the first phase study

The results reported in Table 1, and all subsequent CV results, are based upon "revised" asterisked responses, though these typically differed little from the initial responses.

A comparison of the CV responses for [F3] and [F1] allows a test of sensitivity of CV responses to the magnitude of the risk reduction (i.e. to a 3 in 100,000 and 1 in 100,000 reduction in the annual risk of death). Table 1 shows that, although [F3] affords respondents a three times greater reduction in their risk of death than [F1], aggregate responses do not appear to have been sufficiently sensitive to this difference. A comparison of the means from the whole sample yields a CV [F3]/CV [F1] ratio of 1.41, whereas a ratio closer to 3 might have been expected. This finding is similar to that reported in Viscusi *et al* (1987). Responses do, however, appear to have been more sensitive in the TD group (mean ratio = 1.54), than in the BU (mean ratio = 1.25).

A breakdown of these ratios at the individual level, as shown in Table 2, indicates that 8 respondents in the TD group and 15 in the BU group gave *identical* non-zero CV responses for both magnitudes of risk reduction.[11] This is in spite of the fact that all 23 of these respondents had ranked [F3] higher than [F1] in their initial prioritisation exercise, apparently indicating that they did recognise it to be the larger of the two safety improvements.

Table 1. Revised CV responses from first phase study (£ Sterling)

	TD (N = 40)		BU (N = 41)		Sample (N = 81)*	
	Mean (std error)	Median	Mean (std error)	Median	Mean (std error)	Median
[F1]	111.05 (28.84)	50	85.78 (13.72)	50	98.26 (15.80)	50
[F3]	171.33 (35.39)	100	107.27 (17.58)	67.5	138.90 (19.81)	85
[F3 + P]	234.20 (34.85)	150	170.38 (23.75)	120	201.90 (21.16)	150
[F3 + P + T]	264.74 (34.38)	182.5	189.26 (24.48)	140	226.53 (21.31)	170
[F3 + P + T + M]	267.68 (33.86)	200	191.57 (24.23)	150	229.15 (21.04)	175

*One outlier has been removed from the TD sample (a respondent in the lowest income group category whose first willingness to pay response was £5,000) In addition, one respondent in the TD group failed to give revised CV responses.

Table 2. Individual CV [F3]/CV[F1] ratios

	TD	BU	Sample
= 1	8	15	23
> 1 < = 2	14	17	31
> 2 < = 3	3	2	5
> 3	6	2	8
Missing	9	5	14

1.6 Estimates of the value of statistical life from the first phase study

Computing the value of statistical life (VOSL) from the raw CV responses is relatively straightforward. For example, suppose that an individual indicates that her willingness to pay for a 3 in 100,000 reduction in the risk of death in a road traffic accident during the forthcoming year is £x. That individual's marginal rate of substitution (MRS) of wealth for risk of death in a road traffic accident would then be well-approximated by £x ÷ (3 x 10^{-5}). It can be shown that the VOSL is then given by the population mean of these individual MRS—see, for example, Jones-Lee (1989), Ch. 1. Now suppose that the individual (like the 23 in our sample) is also willing to pay £x for a 1 in 100,000 reduction in the risk of death in a road traffic accident during the forthcoming year. The VOSL estimate based on her CV response for [F1] will then be exactly three times as large as that based on her CV response for [F3].

While the majority of respondents were willing to pay something extra for the larger risk reduction, Table 3 shows that the mean VOSL from the whole sample takes on a value of either £4.6 million, based on the CV responses for [F3] or £9.8 million, based on the CV responses for [F1].Thus, in spite of the fact that respondents had been introduced to key concepts prior to the individual interviews and had also been given every opportunity

Table 3. VOSL estimates from first phase study (£ x 10^6)

	TD		BU		Sample	
	Mean (std error)	Median	Mean (std error)	Median	Mean (std error)	Median
[F1]	11.11 (2.88)	5.0	8.58 (1.37)	5.0	9.83 (1.58)	5.0
[F3]	5.71 (1.18)	3.33	3.58 (0.59)	2.25	4.63 (0.66)	2.83

to revise their responses after having been presented with a summary of their answers to the full sequence of CV questions, it would appear that the VOSL can be more or less arbitrarily inflated or deflated simply by manipulating the magnitude of the risk reduction in the CV question concerned. While this inverse relationship between the VOSL and the magnitude of risk reduction is evident in the results of earlier CV studies[12], its emergence in such pronounced form in a study explicitly designed to give respondents every opportunity to assimilate concepts, and to reflect upon and revise their answers, is particularly worrying.

1.7 Scope and sequencing effects

We turn now to the other issues with which the research was concerned, namely scope and sequencing effects. Sensitivity to the scope of the safety improvement may be tested by looking at how CV responses varied in going from [F3] on its own up to [F3+P+T+M] i.e. Q4 through Q1 in the TD sequence and Q2 through Q5 in the BU sequence. The starkest evidence of insensitivity to scope is afforded by those respondents who gave the same non-zero CV response for one risk reduction as for a more inclusive bundle of reductions of which it was a part (sometimes referred to as "perfect" embedding). These results are set out in Table 4, the first row of which shows the number of respondents who gave the same non-zero amount for [F3] and each of the 3 more inclusive "bundles"[13].

Given the way in which the CV questions were framed, it is clear that while respondents' willingness to pay for reductions in the risk of a fatal road injury was elicited directly, in the case of the non-fatal injuries the *incremental* willingness to pay for reductions in the risk of each non-fatal injury must first be inferred from the raw data. For example, willingness to pay for [P] may be taken as willingness to pay for [F3 + P] *minus* willingness to pay for [F3]. Likewise for [T] and [M]. It is then easy to see that rows 2, 3 and 4 of Table 4 correspond with the number of respondents with implied valuations of

Table 4. Number of respondents completely insensitive to scope

	TD	BU	Sample
[F3] through [F3 + P + T + M]	6	10	16
[F3] = [F3 + P]	8	13	21
[F3 + P] = [F3 + P + T]	20	23	43
[F3 + P + T] = [F3 + P + T + M]	36	29	65

zero for [P], [T] and [M] respectively. Given that more respondents answered in this manner when it came to the more minor injuries [T] and [M] (Table 5 below shows that the median implied willingness to pay for [T] and [M] is zero in both subsamples) it may of course be the case that the consequences of these injuries were considered so trivial that no value was attached to a reduction in their risk.[14] However, one in four respondents also has an implied valuation of zero for [P], a reduction in the risk of an injury which would result in permanent disability and pain. It would therefore appear that the CV responses may not have been sufficiently sensitive to the severity of injury, so that values for the prevention of non-fatal injuries estimated on the basis of incremental willingness to pay will tend to be lower than would have been the case if they had been estimated directly.[15] It should be noted that this also raises the possibility that the VOSL estimates themselves might have been considerably lower had we elected to frame the CV questions in such a way that the valuation sequence was, say, [M], [M + T], [M + T + P], [M + T + P + F1] and [M + T + P + F3] so that values for [F1] and [F3] could only be elicited on the basis of incremental willingness to pay. Were this to be the case then it would consitute a further instance of the susceptibility of CV-based values of safety to manipulation by changes in study design.

Another factor which would render VOSL estimates susceptible to variations in study design would be any tendency for CV responses to be sensitive to the *order* in which the questions were asked. In particular, the results of earlier work had led us to hypothesise that [F1], in appearing first in the BU sequence, would attract substantially higher CV responses than in the TD sequence in which it appeared last. However, contrary to our expectations, the results in Table 1 show the mean CV response for [F1] to be higher in the TD group, although the difference is not statistically significant (Mann-Whitney p value = 0.712). There is, though, a significant difference between the two groups' mean willingness to pay for [F3 + P + T + M], which appeared at the opposite ends of the two sequences (Mann-Whitney p value = 0.048).

Table 5. Implied willingness to pay for reductions in the risks of non-fatal injuries (£ Sterling)

	TD		BU		Sample	
	Mean (std error)	Median	Mean (std error)	Median	Mean (std error)	Median
[P]	67.05 (13.00)	50	63.11 (18.01)	18.5	65.03 (11.10)	34.5
[T]	31.51 (6.46)	0	18.89 (5.11)	0	25.04 (4.13)	0
[M]	4.29 (2.14)	0	13.63 (6.67)	0	9.02 (3.55)	0

2. The second phase study

The findings reported in Section 1 indicate that despite our best efforts, the direct CV questions employed in the first phase study produced results that displayed serious embedding, scope and sequencing effects. Indeed, these effects were so pronounced that it was felt by both the research team and the project sponsors that the type of question used in the first phase study could not be employed with any confidence as a basis for obtaining direct empirical estimates of WTP-based values of safety for use in public policymaking.

While comments by respondents in the individual interviews and feedback meetings suggested a number of possible reasons for the anomalies that were encountered in the first phase study, it appeared that a major contributory factor in the relative failure of the direct CV questions was the difficulty experienced by many respondents in dealing with small reductions in already small probabilities. For example, some respondents seemed to regard one probability reduction of 3 in 100,000 and another of 1 in 100,000 as being essentially the same, since both constitute such small absolute numbers. In view of this and because it has been suggested that people are generally better able to deal with risk information when it is presented in terms of frequencies of occurrence, rather than as probabilities (see, for example, Viscusi *et al* (1991) or Desaigues and Rabl (1995)), it was decided to run a second phase study aimed at eliciting respondents' willingness to pay for the prevention of a pre-specified number of deaths in a given population.

Although previous evidence indicates that presenting risk information in this way *in itself* does not eliminate insensitivity to quantity,[16] it seemed plausible that, when used within the three-stage, in-depth study design outlined above, this alternative approach would increase the sensitivity of responses. In particular, it was anticipated that the inappropriateness of giving the same willingness-to-pay response for the prevention of two different numbers of fatalities would be much more evident and respondents answering in such a manner would be more likely to revise their responses when given the opportunity to do so.

While changing the way in which the risk reduction information is presented may, on the face of it, appear to be a relatively straightforward matter, it raises a whole new set of potential problems for the CV method. The "self only" risk reduction approach adopted in the first phase study had allowed us to deal with a "good" which was strictly private and a payment mechanism—the purchase of a car safety feature—with which respondents could be expected to be reasonably familiar. In contrast, attempting to elicit respondents' willingness to pay for numbers of deaths prevented moves us into the territory of valuing what is essentially a public good, a task which has proved so problematic in the environmental field[17]. On the other hand, this approach does facilitate the use of CV questions involving *household* willingness to pay, which would appear to be a more natural payment concept than that used in the first phase, particularly in relation to certain hazardous contexts, such as automobile safety or domestic fires. It also appeared that many respondents in the first phase study may in any case have been basing their *own* CV response on their *household's* budget.

2.1 Design of the second phase study

The primary objective of the second phase study was thus to test the feasibility of eliciting household willingness to pay for the prevention of a pre-specified number of deaths in a given context. In addition, particular emphasis was placed on the issue of insensitivity with respect to the magnitude of the risk reduction with a view to determining whether embedding effects could be "driven out" of responses. To this end, the design of the second phase study incorporated a number of key changes from that of the first phase. First, more emphasis was placed on the different magnitudes of risk reduction offered by the various safety improvements. Second, the qualitative data suggest that in the first phase some respondents were thinking primarily of the largest amounts that they could *afford* to pay, rather than attempting to assess their maximum *willingness* to pay. In the second phase an attempt was therefore made to focus respondents' attention much more acutely on what the safety improvements were actually *worth* to their household. Third, following the finding of Fischhoff *et al* (1993) that embedding effects are reduced in a simpler study design, no questions concerning non-fatal injuries were included and hence no "bundling" was involved. Thus, the second phase study dealt exclusively with fatal injuries.

In order to provide a direct comparison with the CV[F3]/CV[F1] results from the first phase study, respondents were asked about an improvement which would prevent a number of deaths on the roads in the area in which they live [R1] and another which would prevent three times that number [R3]. In order to test the sensitivity of CV responses with respect to the context in which death occurs, respondents were also asked about household willingness to pay for a domestic fire safety improvement [D1] which would prevent the same number of fatalities as the smaller of the two road safety improvements [R1]. Two variants of the questionnaire were produced—A and F— which differed only with respect to the detailed wording of the CV questions (see below for details). Thus, the study design also allowed us to test the sensitivity of CV responses to framing effects.

A total of 52 respondents—again, broadly representative in terms of age, gender and household income—were recruited by professional market research agencies on a quota sample basis in the Newcastle, Bangor and York areas and the study followed the same three-stage design as was employed in the first phase.

2.2 CV questions used in the second phase study

All respondents were first asked to imagine that there was a program of road safety improvements affecting the area in which they live, with a population of 1 million people, or approximately 400,000 households. Version A respondents were then told that the safety program was expected to prevent 15 deaths on the roads in the next year and asked to consider what this improvement would be worth to their household *over the next year*. Version F respondents were told that the program was expected to prevent 75 deaths on the roads *over the next five years* and asked to consider what this improvement would be worth *in total* to their household.

Respondents were then presented with a series of three boxes.[18] In the first they were asked to state an amount such that they were sure the safety program would *definitely be worth* at least that much to their household. In the second box they were asked to indicate the amount at which they would start to become uncertain whether or not the safety program was worth that much, and in the third box the amount at which they were sure it would *definitely not be worth* that much to their household. Having answered this question, respondents were then presented with two further CV questions, the first of which involved the prevention of 5 road fatalities for Version A and 25 for Version F, while the second involved the prevention of 5 domestic fire fatalities for Version A and 25 for Version F. In summary, denoting the smaller of the two reductions in the number of road fatalities by [R1], the larger by [R3] and the reduction in the number of domestic fire fatalities by [D1], the actual numbers presented in each version were as follows:

		Roads [R3]	Roads [R1]	Domestic Fire [D1]
Version A	(one year)	15	5	5
Version F	(five years)	75	25	25

As in the first phase study, the questionnaire design afforded respondents the opportunity to revise their initial responses. This time, however, we wished to focus respondents' attention on the relative magnitudes of the two reductions in road fatalities. Hence, immediately after completing the two road CV questions, [R3] and [R1], respondents were presented with their "start to become uncertain" responses and prompted:

> In the past, we've found that some people say that preventing 15/75 deaths on the roads is worth three times as much to them as preventing 5/25 deaths on the roads: but other people don't give this answer. Can you say a bit about why you gave the answers you did?

In this way respondents were explicitly asked to consider, and discuss, the sensitivity of their CV responses to the number of road fatalities prevented (and hence, to the magnitude of the risk reduction). Similarly, respondents were asked to consider their responses to [R1] and [D1] in order to provide qualitative data on the issue of sensitivity to context.

2.3 CV results from the second phase study

The results reported in Table 6 are based on the revised "start to become uncertain" responses and are on an annual equivalent basis (i.e. the raw responses in Version F have been divided by 5).[19]

It had been hypothesised (and indeed hoped) that couching the CV questions in terms of the prevention of a number of fatalities would lead to greater sensitivity to scale in responses than had been present in the first phase study. However, the mean responses given in rows 1 and 2 of Table 6 show that willingness to pay for [R3] was again not

Table 6. CV responses from second phase study (£ Sterling)

	Version A N = 26		Version F N = 26		Sample N = 52	
	Mean (std error)	Median	Mean (std error)	Median	Mean (std error)	Median
[R1]	79.30 (25.75)	30	138.33 (32.70)	90	108.80 (21.02)	55
[R3]	95.83 (25.75)	45	196.14 (47.54)	115	145.14 (27.68)	77.5
[D1]	57.30 (16.50)	20	88.52 (15.51)	87.5	72.89 (11.43)	40

anything like three times as much as for [R1]. Indeed, taking the means from the whole sample, a CV[R3]/CV[R1] ratio of 1.33 indicates that there is, if anything, even *less* sensitivity present here than in the first phase (where the equivalent ratio was 1.41). Just as there was more sensitivity in the TD rather than the BU group in the first phase, Version F respondents appear to have shown more sensitivity than those who answered Version A (with the two groups having ratios of 1.42 and 1.21 respectively). It seems plausible that this may have been due to an "absolute numbers" effect, as the more effective road safety program prevented an additional 50 deaths in version F (over 5 years) compared with 10 in Version A.

A breakdown of the CV [R3]/CV[R1] ratios at the individual level is shown in Table 7, from which it can be seen that 42% of respondents in the second phase (compared with 28% in the first) gave *identical* non-zero willingness-to-pay amounts for both road safety improvements, despite the "worth three times as much" prompt reproduced above.

2.4 Estimates of the value of statistical life from the second phase study

While the household willingness-to-pay responses from the second phase CV questions can be used to compute a VOSL in a number of different (but equivalent) ways, it is most straightforward to proceed by multiplying the mean household willingness to pay by the number of households in the area affected (i.e. 400,000) and then dividing the result by the number of fatalities prevented by the safety program concerned. Thus, in the case of [R3] for Version A the VOSL would be computed as: £(95.83 × 400,000) ÷ 15 = £2.56

Table 7. Individual CV[R3]/CV[R1] ratios

	Version A	Version F	Sample
= 1	13	9	22
> 1 < = 2	6	13	19
> 2 < = 3	6	3	9
> 3	1	1	2

million[20]. Notice that under this computational procedure, VOSLs will again be inversely related to the size of the risk reduction whenever CV responses are insensitive to the number of fatalities prevented.

VOSL estimates based on the CV responses reported in Table 6 are given in Table 8.

From Table 8 it is clear that not only are the VOSL estimates inversely related to the size of the risk reduction, but that these estimates also vary according to the way in which the CV questions were framed. Thus, given that Version F respondents answered CV questions involving risk reductions that were simply five-year counterparts to the one-year reductions in the questions put in Version A, we might have expected the implied VOSLs to be broadly similar for the two versions. However, Table 8 shows that the VOSLs for Version F for each of [R1], [R3] and [D1] are substantially larger than their couterparts for Version A. Mann-Whitney tests confirm that the difference is statistically significant at $p = 0.10$ for all three VOSL estimates. The reasons for these differences remain unclear.

Thus, it appears that even within the context of a study design that avoided any reference to probabilities, as such, and which gave respondents every opportunity to reflect upon and refine their preferences, estimates of the VOSL still appear to be susceptible to being influenced to a substantial degree by a number of theoretically irrelevant factors. Furthermore, when the VOSL estimates that they had collectively generated were fed back to respondents in the course of the third-stage follow-up discussions, it was widely felt that the figures concerned were excessive[21].

Finally, on a somewhat more positive note, Table 8 also shows that the mean VOSL for domestic fires is 0.67 times the mean VOSL for roads based on the [R1] responses, indicating that respondents were, on average, willing to pay less to prevent a fatality in a domestic fire than on the roads. Respondents' comments suggested that this differential valuation reflected a) respondents' perception that their own household's domestic fire risks were very low by comparison with the road risks faced by members of their household[22] and b) the view that domestic fire risks are very much more under the potential victim's own control than in the case of road risks, and very much more his or her own responsibility.[23] The domestic fire/roads VOSL differential was confirmed by responses to the matching questions in which 35 of the 52 respondents indicated a preference for a safety program that would prevent 25 road deaths over a program that would prevent 25 domestic fire deaths, where the two programs have identical costs and would be effective

Table 8. VOSL estimates from second phase study (£ x 10^6)

	Version A		Version F		Sample	
	Mean (std error)	Median	Mean (std error)	Median	Mean (std error)	Median
[R1]	6.34 (2.06)	2.40	11.07 (2.62)	7.20	8.70 (1.68)	4.40
[R3]	2.56 (0.69)	1.20	5.23 (1.27)	3.07	3.87 (0.74)	2.07
[D1]	4.58 (1.32)	1.60	7.08 (1.24)	7.00	5.83 (0.91)	3.20

over the same time period. The matching responses therefore suggest that the VOSL *ought* to differ between the two contexts and the responses to the CV questions suggest that the CV approach *is* capable of capturing this differential.

3. Concluding comments

Far from attenuating the pronounced embedding effects encountered in the first phase study, the second phase CV question format, if anything, appears to have exacerbated these effects, with 42% of respondents giving identical non-zero CV responses for [R3] and [R1], compared with the 28% of first phase respondents who gave identical non-zero CV responses for [F3] and [F1]. The fact that in the first phase a further 38% of respondents gave WTP responses for [F3] that were between only one and two times the WTP responses for [F1]—with the corresponding proportion of second phase responses being 37%—merely adds to the already bleak picture. Thus, while conventional economic theory would allow (and indeed, on plausible assumptions, would predict) that the [F3] and [R3] responses would not be as much as three times their [F1] and [R1] counterparts, it does not allow the possibility that these responses will be *identical* if more safety is preferred to less, and would require an implausibly high degree of curvature in the underlying valuation function to accommodate [F3] and [R3] responses that were only between one and two times the [F1] and [R1] figures. Nor is standard theory capable of explaining the marked sequencing and framing effects that are apparent in the results reported above.

Whatever thought processes and strategies respondents brought to bear in answering the CV questions in the two phases, it would therefore appear that in an uncomfortably large proportion of cases these were *not* of the type presupposed by the theory usually taken to underpin the WTP approach to the valuation of safety. One can therefore have little confidence that the VOSL estimates that emerged from these studies can be used as a reliable basis for public policymaking.

While the qualitative data suggest that several factors may have been at work in producing the aberrant response patterns described above, a fairly common feature of many respondents' thinking appears to have run somewhat as follows. First, any safety improvement is seen as a "good thing", with the precise magnitude of the risk reduction being treated as of only secondary importance (and in some cases, no importance at all) as far as CV responses are concerned. In seeking to decide how much this safety "good thing" is worth, in the absence of any significant experience of explicit money/risk trade-offs, many respondents then report an amount which, if foregone, would not seriously disrupt their normal expenditure and savings patterns, which for many people seems to be a sum in the region of £50–£200 per annum. Alternatively, some respondents may partition their overall expenditure in a rough and ready way into separate "mental accounts", with a "reasonable" figure for the safety account vaguely seen as lying in the £50–£200 per annum region.

This having been said, the *reason* for the insensitivity to the magnitude of the risk reduction seems to have differed somewhat between the first and second phase studies. In

Kluwer Journal
@ats-ss5/data11/kluwer/journals/risk/v17n1art1 COMPOSED: 09/10/98 3:57 pm. PG.POS. 16 SESSION: 3

the first phase study, in which safety improvements were presented as very small reductions in already small probabilities, many respondents were clearly unable (or at least unwilling) to discriminate between two different reductions in the probability of death, viewing both as being just "very small numbers". By contrast, in the second phase study, in which safety improvements were presented as the prevention of a particular number of fatalities in a given population, it seems that some respondents saw the "good thing" as the prevention of death *per se*, with the precise *number* of deaths prevented being more or less irrelevant to a decision concerning how much to contribute to the provision of the safety improvement. As such, the sort of thought process which appears to have given rise to insensitivity to the magnitude of the risk reduction in the second phase study seems to have had something in common with that discussed in Kahneman and Knetsch (1992) in which it is hypothesized that in many cases, reported willingness to contribute to the provision of a public good may be largely a reflection of a desire to acquire a sense of "moral satisfaction" from a voluntary contribution to the provision of the good, rather than a concern with the actual *quantity* of the good itself.

If this account of the thought processes underpinning the typical first and second phase CV response patterns is anywhere near the mark, then it suggests that if one is to get at valid and reliable estimates of WTP-based values of safety using the CV approach then it will be necessary to proceed in a less direst, more highly structured way, breaking down the money/risk trade-off into a number of less daunting, more familiar and more manageable steps. More specifically, it would seem that the attempt to forge a link between the money and health impairment dimensions should be made at the level of a less awesome and emotive injury than death. Second, in seeking to establish the trade-off between money and health impairment it would be desirable to deal exclusively with safety improvements that take the form of private goods and, in addition, if possible to avoid any reference to probabilities or relative frequencies.

With this in mind, work on the CV approach subsequent to the first two phases has been aimed at testing the feasibility, reliability and validity of a multi-stage approach that involves "chaining together" responses to CV and standard gamble (SG) questions. The findings of this work were generally encouraging and the project sponsors therefore agreed that the research team should proceed to a main study based on the CV/SG "chained" approach. The main study was carried out during October and November of 1997 and its findings, which make for somewhat happier reading than those reported here, are the subject of Part 2 of this two-part series of articles.

Appendix: Road Accident Injury Description Cards

F
Immediate unconsciousness followed shortly by death.

P
In hospital

- Several weeks, possibly several months
- Moderate to severe pain

After hospital

- Pain/discomfort for the rest of your life
- Restrictions, possibly substantial, to leisure and work activities for the rest of your life

T
In hospital

- Several days, possibly several weeks
- Slight to moderate pain

After hospital

- Some pain/discomfort for several weeks
- Some restrictions to work and/or leisure activities for several weeks/months
- After 3 months to 3 years return to normal health with no permanent disability

M

- No overnight stay in hospital although may be seen as an outpatient
- Cuts and bruises causing some discomfort

After effects

- Minor discomfort/inconvenience for several days
- Fully recovered after several days—no continuing discomfort

Acknowledgments

The research reported in this article was commissioned by the UK Health & Safety Executive; the Department of the Environment, Transport and the Regions; the Home Office and the Treasury. However, the opinions expressed in the article are solely the responsibility of the authors and do not necessarily reflect the views of the project's sponsors. We are grateful to Jonathan Baron, Baruch Fischhoff, George Loewenstein, Paul Slovic, and Kip Viscusi, as well as members of the project Steering Committee, for very helpful comments and suggestions.

Notes

* Because a large number of people contributed to different aspects of the research reported here, we have adopted the convention of listing authors in alphabetical order. Sadly, Jane Beattie died in March 1997. Correspondence regarding this paper should be addressed to one of the four researchers—Judith Covey, Michael Jones-Lee, Graham Loomes and Angela Robinson—who had principal responsibility for the collection and analysis of the quantitative data generated by the research reported in this article.

1. See HM Treasury (1997) and US Department of Transportation, Federal Aviation Administration (1995).
2. See, for example, Jones-Lee (1989), Viscusi (1993) and Beattie *et al* (1998).
3. See Jones-Lee *et al* (1995).
4. Interestingly, the 50% premium for the value of preventing an Underground fatality was entirely the result of "contextual" considerations such as voluntariness, control and responsibility and owed nothing whatsoever to the possibility of large-scale "catastrophic" accidents on the Underground.
5. This consortium comprised the Health and Safety Executive, the Department of the Environment, Transport and the Regions, the Home Office and the Treasury.
6. The UK Department of Transport (DoT) first elected to adopt the WTP approach to the valuation of safety in 1988, setting the value for preventing a road fatality at £500,000 in 1987 prices. Since then, this figure has been updated for inflation and growth of real GDP per capita and now stands at some £903,000 in 1997 prices. The £500,000 figure adopted in 1988 was not the result of any single study but was, rather, a "consensus" estimate arrived at following a comprehensive review of the then existing empirical literature related to the WTP approach, together with extensive consultation with experts in the field—see Dalvi (1988) and Department of Transport (1988). There is therefore a substantial judgemental element in the value for the prevention of a road fatality currently used by the DETR (which now subsumes the DoT)—hence the DETR's decision to commission further empirical work on the WTP-based roads VOSL.
7. This is esentially because computation of individual marginal rates of substitution of wealth for risk—which form the basis for estimating WTP-based values of safety—involves dividing CV responses by the risk reduction concerned. Clearly, if the risk reduction is, say, 1 in a million then each £1 error in an individual's CV response converts to a £1 million error in the corresponding marginal rate of substitution.
8. For a fuller account of these and other anomalies in responses to CV questions in the safety context, see for example Dubourg *et al* (1997), and for a discussion of similar effects in responses to questions aimed at valuing environmental goods, see for example Bjornstad and Kahn (1996).
9. All respondents were provided with a sheet giving the injury descriptions and their associated risks (but not the risk showcards) to take home with them.
10. In this way we aimed to obtain information concerning respondents' "personal confidence intervals" for the monetary sums concerned, as well as their "best" estimates of their willingness to pay (i.e. their asterisked responses). In addition to our desire to estimate personal confidence intervals, there were several other reasons for employing the payment card format in the CV question, rather than the dichotomous choice (DC) format recommeded by the NOAA Panel—see Arrow *et al* (1993). In the first place, the limitations on sample size imposed by the intensive, three-stage study design meant that the DC approach would not have been workable. Second, DC questions would not have permitted the sharp, within-subject, tests of consistency that we wished to conduct. And finally, there is growing evidence that the particular set of "take it or leave it" prices chosen to be presented to the various subsamples of respondents in a DC study may have a substantial impact on the mean willingness to pay inferred from the study—see for example Desvousges *et al* (1996, p 141) and Schulze *et al* (1996, pp 106–110).
11. In addition, 8 respondents gave a zero willingness to pay for both [F1] and [F3] while a further 6 gave a zero response for [F1] but a non-zero response for [F3]. Clearly CV[F3]/CV[F1] ratios cannot be computed for any of these 14 cases.
12. For within-study evidence of the effect, see for example Jones-Lee *et al* (1985) or Jones-Lee *et al* (1995). To the best of our knowledge the only attempt to explore the relationship between estimates of the VOSL and the magnitude of the risk reduction in CV questions on a *between-study* basis is that reported in Beattie *et al* (1998) where the inverse relationship between the VOSL (taken as the sample mean MRS of wealth

for risk for each study) and the magnitude of the risk reduction, δ r, is found to be highly significant with $\log_{10}$ VOSL = 4.14 − 0.52 $\log_{10}$ δ r.

13. One respondent in the BU group gave zero responses for all of the CV questions.
14. The qualitative data tend to confirm this, particularly for injury M. However, some respondents did indicate that they would have been prepared to pay some non-zero amount if reductions in the risk of each non-fatal injury had been valued separately.
15. Further evidence of this is provided by the fact that 26 respondents had an implied value for [P] that was lower than their stated willingness to pay for [F1] despite having indicated a strict preference for [P] over [F1] in the prioritization exercise.
16. See, for example, Desaigues and Rabl (1995).
17. However, it is important to appreciate that whereas many environmental studies seek to elicit *non-use* values, in the approach employed in the second developmental phase respondents would themselves be members of the population at risk and would therefore benefit directly from the risk reduction, were it to be effected.
18. Given that they might experience some difficulty in thinking about the value of a five-year program, Version F respondents were told that they might want to begin by stating a "per year" amount and then multiplying by 5. Accordingly, the questionnaires presented to this group contained boxes for both "annual" and "five year" amounts. Of course, strictly speaking, the conversion of a "per year" amount into an overall present value amount ought to have involved some discounting, but it was felt that any reference to discounting would add a further serious complication to what for most respondents was an already difficult exercise and would in any case impart an element of spurious precision to calculations aimed principally at delineating broad orders of magnitude.
19. Given that it had been suggested to Version F respondents that they might wish to calculate their "five year" amount by first stating a "per year" amount and then multiplying the latter by 5, it was felt that it would be inappropriate to employ any discounting in the calculation of annual equivalent amounts for the Version F responses. See footnote 18.
20. Alternatively, (but equivalently), the VOSL could be computed by dividing mean household willingness to pay by mean household size (giving willingness to pay per household member) and dividing the result by the reduction in individual risk afforded by the safety program concerned. In the case of [R3] for Group A, under this procedure the VOSL would be given by: £ (95.83 ÷ 2.5) ÷ (15 x 10^{-6}) = £2.56 million.
21. In order to put the VOSL figures into perspective, respondents were given some idea of, for example, the approximate number of secondary school teacher—years or hip replacement operations that could be financed by the sums concerned.
22. This explanation for the differential valuation is very much in line with the finding reported in M^{c}Daniels *et al* (1992) and noted earlier.
23. As already noted, considerations of control and responsibility were among the main "contextual" factors which gave rise to the Jones-Lee and Loomes (1995) finding that the value of preventing an Underground fatality stood at a 50% premium in relation to the corresponding roads value.

References

Arrow, K., Solow, R., Portney, P., Leamer, E., Radner, R., and Schuman, H. (1993) Report of the NOAA Panel on Contingent Valuation. *Federal Register*, 58, 4602–14.

Beattie, J., Chilton, S., Cookson, R., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Robinson, A. and Spencer, A. (1998) *Valuing Health & Safety Controls: A Literature Review*, London, Health and Safety Executive.

Bjornstad, D.J. and Kahn, J.R. (eds) (1996) *The Contingent Valuation of Environmental Resources*, Cheltenham, UK and Brookfield, US, Edward Elgar.

Dalvi, M.Q. (1988) *The Value of Life and Safety: A Search for a Consensus Estimate*, London Department of Transport.

Department of Transport (1988) *Valuation of Road Accident Fatalities*, London, Department of Transport.
Desaigues, B., and Rabl, A. (1995) "Reference Values for Human Life: An Econometric Analysis of a Contingent Valuation in France" in Schwab Christe, N.G., and Soguel, N.C. (eds) *Contingent Valuation, Transport Safety and the Value of Life*, Boston, Kluwer.
Desvousges, W.H. Hudson, S.P., and Ruby, M.C. (1996) "Evaluating CV Performance: Separating the Light from the Heat" In Bjornstad, D.J. and Kahn, J.R. (eds) *The Contingent Valuation of Environmental Resources*, Cheltenham, Edward Elgar.
Dubourg, W.R. Jones-Lee, M.W., and Loomes, G. (1997) "Imprecise Preferences and Survey Design in Contingent Valuation", *Economica*, 64, 681–702.
Fischhoff, B., Jacobs Quadrel, M., Kamlet, M., Loewenstein, G., Dawes, R., Fischbeck, P., Klepper, S., Leland, J. and Stroh, P. (1993) "Embedding Effects: Stimulus Representation and Response Mode", *Journal of Risk and Uncertainty*, 6, 211–234.
H.M. Treasury (1997) *Appraisal and Evaluation in Central Government ("The Green Book")*, London, the Stationery Office.
Jones-Lee, M.W. (1989) *The Economics of Safety and Physical Risk*, Oxford, Basil Blackwell.
Jones-Lee, M.W., Hammerton, M. and Philips, P.R. (1985) "The Value of Safety: Results of a National Sample Survey", *Economic Journal*, 95, 49–72.
Jones-Lee, M.W., Loomes, G., and Philips, P.R. (1995) "Valuing the Prevention of Non-Fatal Road Injuries: Contingent Valuation vs Standard Gambles", *Oxford Economic Papers*, 47, 676–695.
Jones-Lee, M.W., and Loomes, G. (1995) "Scale and Context Effects in the Valuation of Transport Safety", *Journal of Risk and Uncertainty*, 11, 183–203.
Kahneman, D., Knetsch, J.L. (1992) "Valuing Public Goods: The Purchase of Moral Satisfaction", *Journal of Environmental Economics and Management*, 22, 57–70.
Kahneman, D., Slovic, P., and Tversky, A. (eds) (1982) *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge, Cambridge University Press.
M^{c}Daniels, T.L., Kamlet, M.S. and Fischer, G.W. (1992) "Risk Perception and the Value of Safety", *Risk Analysis*, 12, 495–503.
Mendeloff, J. and Kaplan, R.M. (1990) "Are Twenty-fold Differences in 'Lifesaving' Costs Justified? : A Psychometric Study of the Relative Value Placed on Preventing Deaths from Programs Addressing Different Hazards," in Cox, L.A. Jr. and Ricci, D.F. (eds) *New Risks,* New York, Plenum Press.
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, 75–90.
Schulze, W., McClelland, G., Waldman, D., and Lazo, J. (1996) "Sources of Bias in Contingent Valuation" In Bjornstad, D.J. and Kahn, J.R. (eds) *The Contingent Valuation of Environmental Resources*, Cheltenham, Edward Elgar.
U.S. Department of Transportation, Federal Aviation Administration (1995) *Treatment of the Values of Life and Injury in Economic Analysis*, APO Bulletin APO-95-1, Washington, Federal Aviation Adminsitration.
Viscusi, W.K., Magat, W.A., and Huber, J. (1987) "An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks" *RAND Journal of Economics*, 18, 465–479.
Viscusi, W.K., Magat, W.A., and Huber, J. (1991) "Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-Offs for Chronic Bronchitis", *Journal of Environmental Economics and Management*, 21, 32–51.
Viscusi, W.K. (1993) "The Value of Risks to Life and Health" *Journal of Economic Literature*, 31, 1912–1946.