J Risk Uncertainty (2006) 33:165–182
DOI 10.1007/s11166-006-0483-z

Dread risks

Susan Chilton · Michael Jones-Lee · Francis Kiraly · Hugh Metcalf · Wei Pang

Abstract It is a well-established fact that many people view the prospect of premature death by some causes with considerably more disquiet or "dread" than death by other causes. It is equally clear that for most people their personal risk of death by a given cause is also a matter of serious concern. This article reports the findings of a study aimed at estimating the effects of dread and personal risk of death by a specific cause on the willingness-to-pay based Value of Statistical Life (VSL) for that cause.

Keywords Dread · Willingness to pay · Value of Statistical Life

JEL Classification J17

It is by now well-known that people typically view the prospect of premature death by some causes with considerably more dread than they do for other causes—see for example Slovic et al. (1981), Thomas (1981), Mendeloff and Kaplan (1990), McDaniels et al. (1992), Savage (1993), Tolley et al. (1995) and Jones-Lee and Loomes (1995).

S. Chilton · M. Jones-Lee (✉) · F. Kiraly · H. Metcalf · W. Pang
University of Newcastle, Ridley Building, Claremont Road, Newcastle NE1 7RU, United Kingdom
e-mail: michael.jones-lee@newcastle.ac.uk

S. Chilton
e-mail: s.m.chilton@newcastle.ac.uk

F. Kiraly
e-mail: francis.kiraly@newcastle.ac.uk

H. Metcalf
e-mail: h.r.t.metcalf@newcastle.ac.uk

W. Pang
e-mail: wei.pang@newcastle.ac.uk

Sunstein (1997) argues convincingly that, in certain circumstances, a "bad death" premium may be appropriate when valuing avoidance of certain types of death. These circumstances include, amongst others, those types of death that are particularly abhorrent or frightening, lack "voluntariness" and control, or induce feelings of empathy with victims.

Set against this is the possibility that personal baseline risk for the cause of death concerned (the current level of risk faced by the individual prior to any risk reduction policy implementation) also matters to people when forming their risk preferences over alternative options—see, for example, Viscusi (1979) or Covey (2001). Such a concern could interact in some way with people's feelings of dread towards the different risks. This might to some extent explain the rather counter-intuitive finding of Chilton et al. (2002), that the willingness to pay-based Value of Statistical Life (VSL) for rail accidents and fires in public places—arguably fairly highly dreaded under the above criteria—in fact stood at a *discount* relative to the road accidents VSL.[1] Clearly, this could well reflect the fact that dread effects for rail accidents and fires in public places are effectively offset by their low baseline risks relative to road accidents.

Before proceeding further, we believe that it is important to be clear about the precise sense in which the terms "dread" and "risk" are being used in the present context. Clearly, an individual's overall attitude—including fear and anxiety—concerning the possibility that he or she may die prematurely by some particular cause will depend on a variety of factors. In addition to the individual's own personal characteristics and circumstances, these will typically include his or her anticipation of the degree of pain and suffering that will precede death; the degree of voluntariness and control associated with the particular cause of death; the time at which death might occur; responsibility; blame, and so on, as well as the magnitude of the risk (i.e. probability) of the event occurring. In what follows, "dread" will be used as a "catchall" term that applies to all of these factors with the exception of the risk (i.e. probability) of death *per se.*

The way in which dread and personal baseline risk thus defined affect peoples' marginal rates of substitution (MRS) between wealth and risk of death has profound implications for policy making with respect to the VSL for any specific cause of death. In particular, a version of an argument presented in Jones-Lee (1976) indicates that, under reasonable assumptions, an individual's MRS between wealth and risk of death by a given cause will, *ceteris paribus*, be a non-decreasing function of the individual's baseline risk of death by that cause and that it will indeed typically be a *strictly increasing* function. Given that for many everyday causes of premature death, dread and baseline risk may be negatively correlated, it is clear therefore that they could tend to have opposite effects on an individual's MRS of wealth for risk though theory, as such, gives no indication of the *extent* to which one effect can be expected to offset the other. The answer to this question is therefore essentially empirical and may well differ from one case to another.

In the light of these considerations the UK Health and Safety Executive (HSE) commissioned an empirical study to investigate these issues. This required the development of a dedicated methodology based on the premise that both dread and cause-specific baseline risk may matter to people when considering wealth-risk tradeoffs and, further, that the two

[1] The Value of Statistical Life is defined essentially as the aggregate willingness to pay for small individual risk reductions which, taken over the affected group of people, will reduce the expected number of fatalities during a forthcoming period by precisely one. Under appropriate assumptions, this aggregate willingness to pay can be shown to be equal to the arithmetic mean of individual marginal rates of substitution of wealth for risk of death for the group concerned—see for example, Jones-Lee (1976). An alternative terminology—increasingly used by Government and related agencies in the UK—is the "Value of Preventing a Statistical Fatality (VPF)," though it should be stressed that this is *precisely synonymous* with the VSL.

Table 1 Baseline average annual risks

Hazard	Baseline average annual risk
Automobile driver/passenger	1400 in 50 million
Rail*	40 in 50 million
Domestic Fire	400 in 50 million
Fire in Public Place	30 in 50 million
Hazardous Production Plant	250 in 50 million
Pedestrian	800 in 50 million
Murder	250 in 50 million
Drowning	100 in 50 million
Accident in the Home	2000 in 50 million

*Excluding trespassers and suicides.

components are in some way separable. Given the novelty of the approach, a number of validity tests were also built into the investigation. In addition, it should be noted at this stage that in the course of our study it became clear that the methodology uncovered some rather unexpected choice heuristics employed by a number of respondents. This necessitated a complementary theoretical investigation to accommodate such behaviour within our empirical framework, which is based on an extension of the risk-risk methodology pioneered by Viscusi et al. (1991). The results of this are considered later in the paper.

In what follows we report the development and application of our protocol designed to disaggregate an individual's marginal rate of substitution of wealth for risk of death into what appear to be its two main driving factors. We find that while substantial dread elements are indeed present in certain types of death, these appear in some cases to be cancelled out to a large extent by low baseline risks. Nevertheless, this does not always hold, implying support for the observation that the inclusion of dread premia in a cause-specific VSL can only be justified on a case-by-case basis.

The causes of premature death considered in the paper are automobile driver/passenger deaths, pedestrian accidents, accidents in the home, fires in public places, domestic fires, drowning, rail accidents, hazardous production plant accidents and murder.[2] Notice that these are all causes that result in instant (or near-instant) death. Causes that result in protracted periods of pain or suffering prior to eventual death, such as heart disease or lung cancer, are not considered in this study as they are characterised by *different and varying* protracted periods of pain and suffering prior to eventual death. As such they are not suited to our methodology, which focuses on the interaction between dread and baseline risk effects, *ceteris paribus.* It is also important to stress the fact that having specified a given cause of death, no further information was provided to respondents concerning the precise circumstances that would be associated with premature death by that cause (e.g. whether a rail accident would involve a head-on collision or a derailment etc.). This was done in order to ensure that, as far as possible, respondents imputed their own personal perceptions and attitudes to the causes of death concerned.

The baseline average annual risks of premature death in the UK for each of the causes considered as shown in Table 1.

Risks were expressed with respect to a denominator of 50 million for two reasons, namely: a) because for some of the causes of death considered in the study, the baseline risk is so small

[2] Aircraft accidents were not included because the baseline risk for such accidents is minuscule by comparison with other causes and this had led to serious comparability problems in piloting.

that with risk expressed in the form of say, x in 100,000 per annum, it would be necessary to set x as a very small fraction. For fires in public places, for example, in the UK the baseline average annual risk of death is about 0.06 in 100,000. Small fractions might prove to be confusing for some respondents in a sample survey; and (b) because the total population of England and Wales is about 50 million, a fact of which respondents were reminded prior to answering the risk-risk questions.

While respondents were presented with these baseline average annual risks for each of the various causes of premature death considered in the study, it is important to appreciate that they were also asked to indicate whether they regarded themselves as being below, equal to or above average risk for the cause concerned and to give an indication of the extent (if any) of their deviation from average. For each respondent subsequent questions were then based on these "personalised" levels of exposure.

The outline of the paper is as follows. Section 1 provides the theoretical background to the methodology, while Section 2 describes its implementation. Sections 3 and 4 consider the general nature of the broad response patterns that emerged in the study, while Section 5 reports the detailed empirical results. Section 6 then considers the robustness of the estimation procedure and empirical results, while the relative impact of dread effects and baseline risks on the VSL are tentatively examined through an econometric exercise in Section 7. Section 8 concludes.

1 The "risk-risk" methodology

We extend Viscusi et al.'s (1991) "risk-risk" methodology to account for the potential separability of baseline risk and dread effects. Thus, consider an individual who is indifferent between a change δp in the annual risk of death by cause A and a change δq in the annual risk of death by cause B. Arguably, the ratio $\frac{\delta p}{\delta q}$ then reflects *precisely two* factors, namely: (a) any differential in the degree of dread (widely construed) that the individual associates with each of the two causes, and (b) any difference between the baseline risk that the individual faces for each of the two causes. More specifically, it seems reasonable to suppose that $\frac{\delta p}{\delta q}$ will tend to be larger (a) the more death by cause B is dreaded relative to death by cause A, and (b) the larger the individual's baseline risk of death by cause B relative to his/her baseline risk of death by cause A.

Now suppose that before identifying the indifference risk changes δp and δq for death by each of the two clearly identified causes of death, the question is put in a "contextless" format *without* the causes being identified, but with the baseline risks clearly specified. Let us denote the individual's indifference risk changes for this contextless risk-risk question by $\delta \hat{p}$ and $\delta \hat{q}$ respectively. In this case, it would appear that the ratio $\frac{\delta \hat{p}}{\delta \hat{q}}$ can reflect *only one* factor, namely the individual's attitude to any difference between the baseline risks for the two causes. If the question is then subsequently put in a "contextual" format with the baseline risks *and* causes of death clearly identified and the individual's indifference risk-change ratio now becomes $\frac{\delta p}{\delta q}$, then it seems reasonable to suppose that any difference between the contextual ratio $\frac{\delta p}{\delta q}$ and the contextless ratio $\frac{\delta \hat{p}}{\delta \hat{q}}$ must be attributable *exclusively to the relative dread* that the individual associates with each of the two causes with, in particular $\frac{\delta p}{\delta q} > \frac{\delta \hat{p}}{\delta \hat{q}}$ indicating that cause B is more dreaded than cause A and vice versa. While it is not essential to the argument that follows, it seems most natural—and, indeed, is analytically most straightforward—to assume that the individual's dread effect for context B relative to that for context A, denoted by D_{BA}, is essentially given by the ratio of $\frac{\delta p}{\delta q}$ relative to $\frac{\delta \hat{p}}{\delta \hat{q}}$ that is

$$D_{\mathrm{BA}} = \frac{\delta p/\delta q}{\delta \hat{p}/\delta \hat{q}} \tag{1}$$

with D_{BA} being larger, the more death by cause B is dreaded relative to death by cause A.

2 Implementation of the dread risk elicitation protocol

The protocol was developed and refined over a three-month period prior to implementation in order to ensure that in the main study respondents would understand the various questions and to highlight areas where additional practice examples and discussions would still be necessary. In this way it was hoped to minimise any errors arising because of misunderstanding of the task, as opposed to its innate difficulty, so that the use of any unexpected response strategies would be more likely to reflect real preferences rather than random choices.

Respondents in the main study participated in a relatively intensive focus group-like procedure with some open-ended discussion. However, all risk-risk tradeoff questions were answered on a strictly individual basis. Below, we summarise the main procedures/stages within the protocol.

In the first stage, focus group participants were introduced to the general nature of the study and the risk concepts that they would be dealing with. The purpose of this stage was to ensure that, when answering the risk-risk questions, the respondent was aware of the size of the average baseline levels of risk and to give an indication on a questionnaire response sheet whether they regarded themselves as being much lower than average, below average, average, above average or much higher than average risk for the various causes of death to be considered in the study. Thus, for example, a respondent who only rarely travelled by rail might reasonably be expected to regard him/herself as being at much lower than average risk of death in a rail accident. Respondents were then asked to "quantify" their own perceived personal risk for each cause. For example, a respondent who was told that the average annual risk of death in a rail accident was 40 in 50 million and believed that she was "much lower than average" might allocate a number of 5 in 50 million. All subsequent risk-risk tradeoffs in the study were based on these person-specific perceptions of the risk. This process more accurately reflects the actual situation facing the respondent and avoids any errors inadvertently introduced by respondents failing to take experimenter-defined probabilities at face value, a potential problem noted by Viscusi (1989).

Given that understanding and answering "risk-risk" questions is almost certainly a difficult task for anyone not already familiar with this sort of exercise, the focus group protocol then moved on to a "practice and familiarization" session in which respondents were presented with a "contextless" risk-risk question in which the potential causes of death were not identified and in which the baseline risks differed from any of those that had been discussed in the first stage. In this question respondents chose between increasing the annual risk of what was labelled "Accident 1" by 10 in 50 million or, alternatively, increasing the annual risk of "Accident 2" by 10 in 50 million, with the baseline risks for the two causes set at 20 in 50 million and 550 in 50 million respectively.[3] Depending on which of the two risk increases was chosen by the

[3] Risk increases rather than reductions were employed essentially because, at least in whole number terms, there is a lower bound to the magnitude of feasible risk reductions. Thus suppose that a respondent is initially

respondent, that increase in risk was then raised until the respondent swapped over and selected the (lower) risk increase for the other cause of death. There then followed a fairly extensive group discussion of each respondent's thought process and reasons for choosing in the particular way that he or she had in fact done.

Quite apart from its importance as a means of familiarizing respondents with the "risk-risk" question format, this practice session was also intended to remove from respondents' immediate consciousness the actual baseline risks underpinning each of the causes of death to be considered later in the session. More specifically, if the subsequent "contextless" risk-risk questions were to serve their intended purpose, then ideally these questions should be answered without any knowledge of the causes giving rise to the risks concerned. Encouragingly, *all* respondents in the follow-up qualitative study confirmed that the "cause-disassociation" objective had been achieved.

Following the practice and familiarization exercise, the focus group protocol then moved on to the "contextless" phase *per se* in which respondents—on an individual rather than group basis—answered six "risk-risk" questions in which the *nature* of the cause concerned was deliberately *not* identified so that responses would in principle reflect *only* respondents' attitudes to personal baseline risk levels rather than dread. In all cases respondents initially chose between increasing the annual risk of what was labelled "Accident C" by x in 50 million or increasing the annual risk of the other cause (e.g. labelled "Accident *A*," "Accident B" etc.) by x in 50 million with x set at either 10 or 30 for each pairwise choice depending on the magnitude of the actual baseline risk. As already noted, each respondent was provided with quantitative baseline risk information that had been "tailored" to their earlier responses to the question concerning their perceived exposure to the particular cause which (unknown to them in the "contextless" phase) in fact underpinned the baseline risk level concerned. Having selected the cause (say Accident C) which they would prefer to have the increment of x in 50 million added to the specified baseline risk, they were then asked how large the increment in the risk of Accident C would have to be before they would switch to a preference for having an increment of x in 50 million added to the specified baseline risk for the other cause (say Accident B).

The focus group protocol then proceeded to its final "contextual" phase which began with an open-ended discussion of the various different accident causes clearly identified (e.g. rail accident or fire in a public place) followed by a ranking exercise in which each participant ranked the accident causes from "worst" (i.e. the one they most dreaded) to "best" (i.e. the one they least dreaded). More specifically, respondents were asked to imagine that, for whatever reason, they were inevitably going to die by one of the identified causes in the very near future and were then asked to rank the causes in terms of the extent to which they feared or dreaded death by each cause. Quite apart from affording an initial focus on the notion of dread, this exercise provided the basis for a test of convergent validity of the empirical procedure that we used to estimate dread effects for the various causes of death, in that a minimal requirement for such validity would appear to be a close correspondence—*at the*

faced with a choice between a decrease of 10 in 50 million in the annual risk of death by cause *A* and a decrease of 10 in 50 million for cause *B*. Suppose also that cause *B* is dreaded so much more than cause A that the individual's MRS of wealth for risk of death by cause *B* is twenty times that for cause *A*. The individual would therefore express a strict preference for reduction in the risk of death by cause *B* by 10 in 50 million and should only become indifferent if the reduction was brought down to $1/2$ in 50 million, which for many people would be meaningless or at best confusing. By contrast, in realistic terms, there is no corresponding upper bound for risk increases.

level of the individual—between these estimated dread effects and a respondent's ranking of causes in terms of fear or dread.

Following the ranking exercise, participants were then handed a sheet of paper that gave them the average annual baseline risk for each cause as well as their earlier assessment of the risk they personally faced and they were asked to add a few words or a sentence summarising their feelings towards the different causes. Armed with this information, each participant then answered five contextual risk-risk questions in which the causes of the accidents concerned were clearly identified.[4]

3 Responding rationally to risk-risk questions

Consider an individual facing the possibility of death in either of two contexts, A and B, with baseline annual risks of say, 400 in 50 million and 10 in 50 million respectively. Suppose in addition that the individual is offered the choice between increasing the risk of death in context A to 410 in 50 million or, alternatively, increasing the risk in context B to 20 in 50 million (i.e. a choice between an increment of 10 in 50 million to one or other of the two risks). *Prima facie* it is tempting to suppose that if the choice is posed in a "contextless" format so that there are no differential dread effects at work, then the individual will be indifferent between the two alternatives since they appear to involve an initial total risk of death of 410 in 50 million prior to the risk increase which would then rise to an overall risk of 420 in 50 million whichever of the two causes the increment of 10 in 50 million was added to.

Certainly, indifference would be implied if (a) the two risks were taken to be mutually exclusive; (b) no differential dread effects applied; and (c) the individual made the choice in a conventionally rational manner. However, if instead the risks were taken to be *independent* then, as shown below, even a completely rational individual would display a *strict preference* for adding the incremental risk to the *lower baseline*, i.e. context B. But why might one view the two risks as being independent rather than mutually exclusive? We believe that there are in fact *very persuasive reasons* for doing so. To see why, consider the (admittedly somewhat extreme) case of an individual facing the possibility of death during the coming period by one of two causes, the first involving a probability of death of 0.9 and the other a probability of death of 0.1. Is the individual certain to die during the period concerned? Most of us would, one suspects, conclude that the answer is "no" and that the individual could—albeit with very low probability—escape death by either cause. But of course if the risks are mutually exclusive then the probability of death is $0.9 + 0.1 = 1$ and death is a certainty. If, by contrast, the risks are taken to be independent then the individual faces a non-zero survival probability of $(1-0.9)(1-0.1) = 0.09$. In this case, therefore, independence appears to make much more sense than mutual exclusivity. Admittedly, as the size of the risks concerned gets smaller, then the difference between independence and mutual exclusively diminishes, but it would nonetheless appear that in general terms independence is the more plausible assumption.

The sceptical reader might, of course, respond that since one presumably can actually die by only one cause and since independence of two events admits the possibility of the

[4] In fact, the sample was split into two subsamples, one of which considered one group of five accidents with a common risk increment of 10 in 50 million while the other considered a slightly different group with a common risk increment of 30 in 50 million. However, as well as containing a common accident (Murder) by which accidents from two subsets could be compared, two other accident types were also common to both subsets, namely car driver/passenger fatalities and pedestrian fatalities. This allowed an in-built consistency check to be carried out on responses to the same types of tradeoff.

occurrence of both, then from a logical point of view independence is ruled out in this context. But our response to this is simple. In particular, suppose the two "experiments" that may result in death are conducted at different times with, say, experiment A preceding experiment B. If experiment A turns out badly (with, say, probability p) then the person dies by cause A and that is the end of the matter. If, on the other hand, the individual survives experiment A (with probability $1 - p$) then experiment B takes place with its probability (say q) of a bad outcome *unaffected* by the happy outcome of the first experiment. Thus, before either experiment takes place, the probability of surviving experiment A and then subsequently dying in B is given by $(1 - p)q$ so that the overall probability of dying by one or other cause is $p + (1 - p)q$ and the overall probability of surviving *both* experiments is therefore $1 - [p + (1 - p)q] = (1 - p)(1 - q)$. Under these circumstances the assumption of independence therefore makes perfectly good sense. However, as already noted, the effective difference between independence and mutual exclusivity depends on the magnitude of p and q. In particular, with p and q both small then the product pq will be negligible and the numerical difference between independence and mutual exclusivity trivial. On the other hand, if p and/or q are relatively large the difference may be substantial.

But what, then, can we say about rationality? Faced with a choice between adding an increment x to either p or q and treating the experiments as a sequence of independent trials, in the absence of any differential dread effects an individual who was rational in the conventional sense would simply compare overall survival probability $(1 - p - x)(1 - q)$ with overall survival probability $(1 - p)(1 - q - x)$ and would select whichever was the larger of the two. Since $(1 - p - x)(1 - q) > (1 - p)(1 - q - x)$ if and only if $(1 - p)x > (1 - q)x$, it follows that with $x > 0$ the necessary and sufficient condition for the individual to prefer the increment in risk to be added to q rather than p is $q < p$, that is, the individual would strictly prefer to add the increment in risk to the *smaller* of the two baseline risks.

Furthermore, in the absence of any differential dread effects the increment y which, when added to the lower baseline probability q, would render the individual *indifferent* to the addition of x to p would be such that:

$$(1 - p - x)(1 - q) = (1 - p)(1 - q - y) \tag{2}$$

or equivalently:

$$y = \frac{(1 - q)}{(1 - p)}x. \tag{3}$$

Two results follow immediately from Eq. (3). First, though with $q < p$ it is necessarily the case that $y > x$, nonetheless with p, q and x all small it follows that the extent to which y will exceed x will also be small. On the other hand, from Eq. (3) it also follows that as $x \to 1 - p$, so $y \to 1 - q$ or equivalently $q + y \to p + x$, that is, the indifference increment y gets closer to that which would cause the absolute risk of death by cause B, $q + y$, to be equal to the absolute risk of death $p + x$ by cause A.

In the light of these conclusions concerning the possible ways in which a rational individual might be expected to respond to a contextless risk-risk question, we now consider the ways in which our respondents did in fact deal with these questions and the extent to which these actual response patterns can be taken to constitute a satisfactory basis for inferring overall dread effects.

4 The actual risk-risk response patterns

Early in the study it became clear that in answering the contextless risk-risk questions in particular, respondents were, broadly speaking, employing one of three heuristics (or at least variants thereof). Thus again suppose that we have two contexts, A and B, with baseline annual risks of 400 in 50 million and 10 in 50 million respectively. Faced with a choice between an increase of 10 in 50 million in the risk of either A or B, the extreme variants of the three heuristics (and we would stress that these are the *extreme* variants) would then be as follows:

- *Incremental risk equalization* (IRE) would entail indifference between the two increments, so that the initial choice would be made on a random basis. However, as argued above, while indifference would certainly be entailed if the two risks were taken to be mutually exclusive, the latter assumption appears to be somewhat implausible, with independence appearing to make rather more sense and in this case the individual would have a strict preference for the increase in risk to be applied to the lower baseline risk i.e. context B. Nonetheless, with a very large risk denominator (in our case 50 million) indifference would require only a *minute* increase above 10 for the incremental risk for B.
- *Absolute risk equalization* (ARE) would entail that the initial choice would be an increment of 10 in 50 million to the risk of B. Only when the incremental risk of B had risen to 400 in 50 million would the two incremental risks be judged indifferent (i.e. an increment of 10 in 50 million to the risk of A and an increment of 400 in 50 million to the risk of B resulting in absolute risk equalization at 410 in 50 million for both risks).
- *Incremental risk-ratio equalization* (IRRE) would entail that the initial choice would be an increase of 10 in 50 million to the risk of A. Only when the incremental risk of A had risen to 400 in 50 million would the two incremental risks be judged indifferent on the grounds that an increment of 400 in 50 million would double the risk of *A*, just as an increment of 10 in 50 million would double the risk of B.

While these three heuristics are, admittedly, fundamentally different, it is important to appreciate that a not entirely implausible rationale can be offered for versions of each of all three. Thus, as already argued above, incremental risk equalization would follow from the assumption of mutual exclusivity of the two risks and would also be closely approximated given independence provided that the risks concerned were sufficiently small. In the case of absolute risk equalization, the assumption of independence, together with a tendency to ignore or mentally deflate the size of the risk denominator of 50 million would quite reasonably lead to application of a variant of this heuristic. In turn, incremental risk-ratio equalization would tend to follow from an understandable (if strictly misguided) tendency to focus upon the fact that, for example, a 2.5% increase from 400 to 410 is very small relative to a 100% increase from 10 to 20.

Nonetheless, it is vital to appreciate that despite the difference between the three heuristics, the ratio of dread premia entailed by Eq. (1) will in fact be *quite unaffected* by which heuristic is actually used, provided that the respondent employs *the same basic heuristic* in answering both the contextless and contextual risk-risk questions. To see why, suppose for example that context *A* is low-dread while context *B* is high-dread. Further suppose that we have a society of six people of which individuals 1 and 2 are incremental risk equalizers, individuals 3 and 4 are extreme absolute risk equalizers and individuals 5 and 6 are extreme incremental risk-ratio equalizers.

In the contextless case, faced with an initial choice between an increase of 10 in 50 million in the annual risk of either A or B, with the chosen option then having its risk increased up to the point of indifference, given that the other risk increase is held constant at 10 in 50 million, the indifference risk increases would then be:

Individual	*A*	*B*
1	10	10
2	10	10
3	10	400
4	10	400
5	400	10
6	400	10

In turn, suppose that once the contexts are identified, given that the dread factor for B substantially exceeds that for *A*, then the indifference risk increases alter to:

Individual	*A*	*B*
1	20	10
2	20	10
3	10	200
4	10	200
5	800	10
6	800	10

Consider first individuals 1 and 2 (the incremental risk equalizers). For these individuals we have:

$$\frac{\delta\hat{p}}{\delta\hat{q}} = \frac{10/(50 \times 10^6)}{10/(50 \times 10^6)} \tag{4}$$

and

$$\frac{\delta p}{\delta q} = \frac{20/(50 \times 10^6)}{10/(50 \times 10^6)} \tag{5}$$

so that from Eqs. (1), (4) and (5) we have:

$$D_{\mathrm{BA}} = 2 \tag{6}$$

that is for the incremental risk equalizers the dread effect for context *B* relative to that for context *A* is 2.

In turn, for individuals 3 and 4 (the extreme absolute risk equalizers) we have:

$$\frac{\delta\hat{p}}{\delta\hat{q}} = \frac{10/(50 \times 10^6)}{400/(50 \times 10^6)} \tag{7}$$

and

$$\frac{\delta p}{\delta q} = \frac{10/(50 \times 10^6)}{200/(50 \times 10^6)} \tag{8}$$

so that from Eqs. (1), (7) and (8) we have:

$$D_{\mathrm{BA}} = 2. \tag{9}$$

Finally, for individuals 5 and 6 (the extreme incremental risk-ratio equalizers) we have:

$$\frac{\delta \hat{p}}{\delta \hat{q}} = \frac{400/(50 \times 10^6)}{10/(50 \times 10^6)} \tag{10}$$

and

$$\frac{\delta p}{\delta q} = \frac{800/(50 \times 10^6)}{10/(50 \times 10^6)} \tag{11}$$

so that from Eqs. (1), (10) and (11) it follows that

$$D_{\mathrm{BA}} = 2. \tag{12}$$

Clearly, therefore, while the extreme versions of each of the three choice heuristics not surprisingly produce very different indifference incremental risk ratios, the implied relative dread effects are *precisely the same* in all three cases. This strongly suggests that more modest variants of each of the three heuristics would also yield the same result. In the light of this, we feel confident that provided a respondent employs the same basic heuristic in answering both the contextless and contextual risk-risk questions, then estimation of the respondent's relative dread effect for the two causes of death concerned using Eq. (1) will, to all intents and purposes, yield a result that is effectively *independent* of the precise heuristic employed by the respondent. In order to obtain empirical estimates, we have therefore proceeded by computing individual relative dread effects on the basis of Eq. (1).

Having considered the estimation of individual dread effects, the question that then arises is how these individual effects are to be aggregated into an overall sample estimate. The first point to note is that as in most stated-preference empirical exercises, this study produced a small number of extreme outliers which have a *very* marked effect on the untrimmed sample arithmetic mean. For example, in the case of rail relative to pedestrian fatalities, the untrimmed arithmetic mean relative dread effect for our sample of some 140 responses was 14,532:1, whereas with the top and bottom four outliers trimmed out this fell to 185:1. By contrast, the median rail : pedestrian relative dread effect was 3.1:1 while the untrimmed geometric mean was 8.6:1. To the extent (a) that there are bound to be doubts about the reliability of extreme outlier responses and (b) that it is ethically questionable to allow a few individuals to have a very marked effect upon potentially important safety-related policy issues, then there would seem to be a powerful argument in favour of focusing upon the sample median and/or geometric mean responses as the key central tendency indicators of public attitudes towards dread. But particularly in the case of ratios—which is precisely what dread effects (as we have defined them) actually are—there is a further very powerful *a priori* argument in favour of using the geometric mean provided that there are no zero observations as will necessarily be the case in a study such as this.[5] In particular, consider a two-person society in which individual 1's relative dread ratio for two contexts A and B was 2:1 while the other person's ratio was 0.5:1. For the society as a whole, weighing the two individual's interests equally, there would appear to be no grounds whatsoever for treating the overall social dread ratio as being other than 1:1 which is, of course, *precisely* the result given by the geometric mean of the individual responses and *not* by either the arithmetic mean or median, both of which would yield a ratio of 1.25:1. Notice that exactly the same result would emerge even if the relative dread ratios were inverted (i.e. the dread of context B relative to context A rather than *A* relative to *B*). All things considered therefore, there would appear to be a very persuasive practical and ethical case in favour of using the *geometric* mean as the appropriate central tendency measure in a study such as this.[6]

5 Findings of the focus group study

Focus group sessions—which typically involved four participants selected on a quota basis from Newcastle, Norwich and Edinburgh by a professional sample survey organisation—were carried out during November and December 2003 and were moderated by two of the authors of this paper and two trained research assistants. The total sample size was 157, of which 112 participants employed some variant of the absolute risk equalization heuristic in answering the risk-risk questions, 22 some variant of the incremental risk-ratio equalization heuristic and 11 the incremental risk equalization heuristic, the remaining 12 participants having failed to provide a full set of usable answers to the risk-risk questions.

Given that its baseline risk falls in the middle of the range of risks considered in the study—and given that everyone is, in some sense, subject to this risk—murder was treated as the common "comparator" context in the risk-risk questions. Individual dread effects were therefore first computed relative to murder. These individual effects were then normalised with respect to the individual's dread effect for pedestrian accident relative to murder, as pedestrian accident appeared to be the least dreaded of those contexts that were considered by all participants in the study (see footnote 4). The results are reported in Table 2 along with the sample arithmetic mean initial ranking of the contexts from 1 (least dreaded) to 5 (most dreaded). Recall that this ranking exercise was undertaken by respondents on an individual basis *prior* to answering the contextual risk-risk questions. As such, the initial rankings on the one hand, and the estimated dread effects on the other, represent *independently derived* rankings of the same set of preferences. As a result of the normalization, murder does not appear in Table 2 or in the rest of the data analysis. However, focusing on the original "dread" effects relative to murder it transpires that *all* of the other contexts had a geometric mean "dread" effect less than 1, indicating that murder is the most dreaded of all of the contexts considered in the study.

Before proceeding to discuss the study's findings, we believe that it is important to stress the fact that the estimated dread effects reported in Table 2 were derived from risk-risk

[5] Even if a respondent gave a "hard to choose" risk increment response entailing the certainty of death by a given cause, this increment would still necessarily be less than 50×10^6 so that the implied relative dread effect, though very small, would still be non-zero.

[6] In fact there is one other procedure that could sensibly be used in order to arrive at a central tendency measure in a study such as this—see Chilton et al. (2002). However, this procedure is considerably more complicated and less direct and does not produce substantially different results, so that in the interests of expositional simplicity we have elected to focus on the geometric mean in this article.

Table 2 Ranking and "dread effect" relative to pedestrian accident

Accident cause	Rank	Dread effect (Geometric mean)
Set S[1]		
Pedestrian accident	1.00	1.00
Accident in the home	1.00	0.81
Automobile driver/passenger accident	1.22	1.67
Rail accident	1.33	8.65
Fire in a public place	2.32	5.80
Set B[1]		
Pedestrian accident	1.00	1.00
Automobile driver/passenger accident	1.10	1.40
Hazardous production plant accident	1.39	1.53
Drowning	1.53	1.88
Domestic fire	2.00	1.45

1. Set S = a common risk increment of 10 in 50 million; Set B = a common risk increment of 30 in 50 million.

questions in which personal baseline risk had been set (in both the contextless and contextual cases) at levels that reflected the respondent's *own assessment* of the extent of his/her personal exposure to the risk of the particular cause of death concerned. To this extent, we are confident that the dread effects reported in Table 2 really do reflect dread *per se*, rather than the level of personal exposure to risk.

The first and rather encouraging point to note about the results reported in Table 2 is the clear *prima facie* evidence of convergent validity reflected in the relatively close correspondence between the means of respondents' initial ranking of the contexts in terms of dread and the ranking in terms of our estimated "dread" effects. In turn, given that the original "dread" effects relative to murder plainly indicated that the latter was the most dreaded of all contexts, it seems that murder, rail accidents and fires in public places can all be regarded as being clear dread risks in the eyes of the public. In the case of automobile driver/passenger accidents, hazardous production plants, drowning and domestic fires the dread effects are less marked, while the only contexts that show up as having no "dread" effect—at least relative the other causes—are pedestrian accidents and accidents in the home. Reassuringly, it also turns out that the dread effects for automobile driver/passenger in samples S and B are not statistically significantly different from each other.

Given these findings we now turn to a consideration of their validity and reliability.

6 Validity of the estimation procedures

We apply three distinct validity tests to our procedures and/or data. These are considered in turn.

6.1 Theoretical validity

As already argued above in Section 4, fairly plausible processes of interpretation and analysis of the data presented in the questions could quite reasonably lead a respondent to adopt a variant of any one of the three heuristics that appeared to underpin the answers to the risk-risk questions posed in the study (though in the case of incremental risk-ratio equalization it has to be admitted that the rationalization offered was perhaps somewhat less persuasive than for the other two heuristics). Furthermore, at least in the case of the illustrative question

Table 3 Estimated dread effects for the three heuristics

Accident Cause	Dread dffect (Geometric mean)			
	Whole sample ($n = 145$)	IRE ($n = 11$)	ARE ($n = 112$)	IRRE ($n = 22$)
Pedestrian	1.00	1.00	1.00	1.00
Automobile Driver/Passenger	1.53 (1.29)*	2.37	1.01	11.15 (4.8)*

*Four IRRE outliers trimmed out.

and responses discussed above in Section 4, it transpired that while the three heuristics not surprisingly produced very different responses to the risk-risk questions, our proposed procedure for inferring dread effects from the responses to these questions in fact produced *identical* results for each of the three heuristics.

As far as the actual survey results themselves are concerned, had the subsamples of respondents employing each of the three heuristics been sufficiently large, then it would clearly have been appropriate to compute and compare the dread effects estimated from each of the three subsamples. However, in the case of incremental risk-ratio equalizers and—even more so for incremental risk equalizers—the subsample sizes are plainly too small to make this a very meaningful exercise. For example, in the case of hazardous production plant, only *one* of the useable Set B responses was provided by an incremental risk equalizer.

Nonetheless, it was the case that both Set S and Set B respondents answered risk-risk questions for pedestrian as well as automobile driver/passenger accidents, so that it seems not entirely unreasonable to compare the dread effects for automobile driver/passenger relative to pedestrians across the three subsamples. In doing so it should, however, still be borne in mind that the subsample sizes for incremental risk-ratio and incremental risk equalizers are still very small (22 and 11 respectively), so that dread effect estimates for these subsamples are particularly prone to outlier effects. Table 3 therefore presents estimated geometric mean dread effects for automobile driver/passenger relative to pedestrian for each of the three subsamples, both on an untrimmed basis and also with four incremental risk-ratio outliers removed.

Plainly, with the notable exception of the incremental risk-ratio equalization subsample, the results presented in Table 3 are broadly encouraging in that, relative to the rankings and dread effects for, say, rail accidents or fire in a public place reported above in Table 2, the IRE and ARE subsample dread effects both point broadly in the same direction, indicating that while automobile driver/passenger accidents are a little more dreaded than pedestrian accidents, the IRE and ARE dread effects for the automobile case are not of the same order of magnitude as rail or fire in a public place. Most significantly, this result tends to allay concerns that the "absolute risk equalization" heuristic may have biased the study's findings relative to the other heuristic that *prima facie,* seems to sit most comfortably with conventional choice theory, namely IRE. Turning to the IRRE responses, it is clear that outliers are doing a considerable amount of work, with the estimated dread effect being more than halved by removal of 4 (out of 22) responses. Indeed if a further 4 are removed then the figure falls to 1.70. *Prima facie*, therefore, the IRRE responses appear to be somewhat at odds with the rest of the sample, which is perhaps not so surprising given that, as indicated above, the rationale for the IRRE heuristic is rather less persuasive than that which might be taken to underpin the other two heuristics.

All things considered, therefore, it would appear that as far as the theoretical validity of our estimation procedure is concerned the message is generally on the positive side.

6.2 Face validity

As far as face (or "content") validity is concerned, a follow-up qualitative study carried out on thirteen participants indicated a high degree of understanding and careful thought concerning the questions that respondents were asked. Importantly, in 12 out of 13 cases, respondents described a consistent use of their chosen risk heuristic across the contextless and contextual questions, which is of course necessary if the argument underpinning Eq. (1) and our subsequent interpretation of responses are to hold. Further, respondents appeared to give answers that were indeed reflective of their feelings of fear and dread concerning the premature fatality risks in different contexts.[7]

6.3 Convergent validity

It is important to recall that prior to answering the contextual risk-risk questions, respondents were asked to imagine that they were inevitably going to die by one or another of the various causes of accidental death in the very near future and then to rank the causes in terms of fear or dread. Clearly, a comparison of each respondent's direct ranking of the causes in terms of fear or dread on the one hand with the ranking entailed by the dread effects estimated from the respondent's answers to the contextless and contextual risk-risk questions on the other, provides a very direct test of convergent validity at the *individual* level given that, in an ideal world, for each individual respondent the two rankings should be *identical.* In view of this, we carried out a two-tailed Sign Test of Equality of Matched Pairs[8] which provided *no evidence* of a significant difference between individual rankings *even at the* 60% *level*. This we regard as convincing evidence of the convergent validity of our dread effect estimation procedure and the consistency of our respondents.

Given the potentially confounding impact of a few extreme upper-tail outliers, we then ran a regression of the log of individual dread ratios against dummies for all of the accident types, controlling for the effect of the gender, age and the income of each respondent. The anti-logs of the resulting 'dread coefficients' on each of the accident types were insignificantly different to the previously reported geometric means and this adds further support to the convergent validity of our dread effect estimation procedure.

7 Dread effects vs baseline risks

In order to apply our findings at an aggregate (i.e. policy-relevant) level, it would be necessary to have access to estimated VSLs for all of the relevant contexts, along with baseline risk and estimated dread effects taken from the same sample. It would then in principle be possible

[7] The results of the follow-up qualitative study were analyzed by Dr. Rachel Baker and are available from the authors on request.

[8] See Snedecor and Cochran (1989).

to examine the way in which baseline risks, as well as dread effects, impact statistically on the VSLs for the contexts concerned. Unfortunately, however, at the time of writing, to the best of our knowledge such a comprehensive data set does not exist. One might therefore reasonably leave it to the reader to take our estimated dread effects—along with baseline risk data—and compare these informally with those VSL estimates that do exist—albeit derived from different studies/samples. The latter include, in particular, the U.K. roads VSL—see Carthy et al. (1999)—together with pre- and post-Ladbroke Groves estimates of the UK rail VSL, as well as VSL estimates for fires in the home and fires in public places—see Chilton et al. (2002).

However, freely admitting the extremely small sample size and the questionable use of data derived from different samples and at different times, for illustrative purposes we nonetheless have run a simple linear regression of the form:

$$VSL_i = \alpha B_i + \beta D_i + u_i \tag{13}$$

where B_i is the mean baseline risk in context i, D_i is our estimated dread effect reported above in Table 2 and u_i is a random error term. The function has been constrained to pass through the origin given that, in the hypothetical "perfect-world" case of zero dread and zero baseline risk, one might reasonably expect a zero VSL. The results of this regression analysis are reported in Table 4.

Denoting the roads VSL for subsample S by VSL_{RDS} and for subsample B by VSL_{RDB}, the rail VSL by VSL_{RL} etc., the $VSLs$ predicted from our regression analysis are as shown in Table 5.

In turn, the percentage contribution of baseline risk and dread effects to our predicted VSLs are given in Table 6.

A more detailed account of the regression analysis is available from the authors on request.

Table 4 Estimated regression coefficients

	Pre-Ladbroke Grove		Post-Ladbroke Grove	
	Coefficient	*P* value	Coefficient	*P* value
α	3.96×10^{10}	0.035	3.84×10^{10}	0.028
β	150,107.4	0.034	165,740.4	0.019

Table 5 Predicted VSLs

	Pre-Ladbroke Grove	Post-Ladbroke Grove
VSL_{RDS}	£1.3595×10^6	£1.3520×10^6
VSL_{RDB}	£1.3180×10^6	£1.3072×10^6
VSL_{RL}	£1.3301×10^6	£1.4644×10^6
VSL_{PF}	£1.8944×10^6	£1.9843×10^6
VSL_{DF}	£0.5344×10^6	£0.5475×10^6
VSL_{AIH}	£1.7056×10^6	£1.6686×10^6
VSL_{HPP}	£0.4277×10^6	£0.4456×10^6
VSL_{DR}	£0.3614×10^6	£0.3884×10^6
VSL_{PED}	£0.7837×10^6	£0.7737×10^6

Table 6 Contribution of baseline risk and dread effects to predicted VSLs

	Pre-Ladbroke Grove		Post-Ladbroke Grove	
	Baseline Risk	Dread	Baseline Risk	Dread
RDS	82%	18%	80%	20%
RDB	84%	16%	82%	18%
RL	2%	98%	2%	98%
PF	3%	97%	2%	98%
DF	59%	41%	56%	44%
AIH	93%	7%	92%	8%
HPP	46%	54%	43%	57%
DR	22%	78%	20%	80%
PED	81%	19%	79%	21%

8 Conclusion

The results reported above were generated from a new protocol designed specifically to isolate and measure the extent of an individual's dread of premature death by a given cause and, in particular, to quantify the impact of this dread effect on the individual's marginal rate of substitution of wealth for risk of death by that cause. In attempting to do this, it is clearly essential that the protocol concerned should be capable of separating the dread effect from the other factor that past research suggests has a major impact on the individual's MRS of wealth for risk, namely their perception of the baseline level of their own personal exposure to the risk concerned. During the process of development and implementation of the protocol it became clear that some difficult theoretical and empirical issues had to be addressed if the results were to have any validity in and of themselves and, further, if the procedures were to be potentially useful in safety policy making in the future.

In view of this, the protocol was subjected to various validity tests with respect to which it appeared to perform fairly well. However, it has been suggested to us that an unacceptably large number of respondents may well simply not have understood the questions that we put to them—hence the prevalence of variants of the absolute risk equalization heuristic. Our response to this criticism is twofold.

First, as we have already noted, the results of our various validity tests and follow-up qualitative interviews suggest that pervasive confusion on the part of the respondents—and the randomness of the responses which one might expect such confusion to generate—are simply not evident in the quantitative or qualitative findings of our study. Thus, we conclude that our responses in the main reflect consistent (and well-understood) behavioural choices on the part of our respondents. While the majority of these responses perhaps differ to some extent from what might be expected on the basis of conventional choice theory, this is nonetheless the way in which participants in our study *actually chose* between options in a variety of scenarios.

Second, the probability questions that we put to our respondents were as simple and straightforward as it is possible to make such questions. In addition, early exercises in the group protocol helped respondents to understand the questions and the effect on their overall risk of death of alternative response strategies. Thus, to ask first whether a respondent would prefer to have an unavoidable increment of 10 in 50 million added to the baseline annual risk of death in context A or in context B and then if, say, context A is chosen, to iterate the increment in context A upward until the respondent regards it as being equally as bad as the incremental risk of 10 in 50 million in context B, does not seem to us to be excessively complicated, even if somewhat demanding. Data from the qualitative study supports this conclusion. Indeed, if such questions were too difficult for our respondents, then we are bound to ask how *any confidence whatsoever* can be placed in the findings of earlier risk-risk studies or, for that matter, *any* stated-preference studies in the safety field. While we accept that by their very nature, studies of this type cannot be expected to produce high-precision "point" estimates, we nonetheless firmly believe that they are capable of giving a pretty clear and reliable indication of the broad order of magnitude of preference-based costs and values that are an absolute prerequisite for responsible public-sector allocative and regulatory decision making in a free society.

Acknowledgments The research reported in this article was commissioned by the UK Health & Safety Executive. However, the opinions expressed in the article are solely the responsibility of the authors and do not necessarily reflect the views of the research project's sponsors. We are grateful to Judith Covey, Graham Loomes, Angela Robinson and Anne Spencer, as well as members of the Project Steering Committee for very helpful comments and suggestions. Finally, a referee and the editor have both made a number of particularly insightful suggestions for revisions to and extensions of the argument developed in the paper. We are extremely appreciative of this input though again, of course, we bear total responsibility for our argument and conclusions.

References

Camerer, Colin. (1995). "Individual Decision Making," in Kagel, J.H. and Ruth, A.E. (eds.), *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

Carthy, Trevor, Susan Chilton, Judith Covey, Lorraine Hopkins, Michael Jones-Lee, Graham Loomes, Nick Pidgeon, and Anne Spencer. (1999). "On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 2 – the CV/SG 'Chained' Approach," *Journal of Risk and Uncertainty* 17(1), 187–213.

Chilton, Susan, Judith Covey, Lorraine Hopkins, Michael Jones-Lee, Graham Loomes, Nick Pidgeon, and Anne Spencer. (2002). "Public Perceptions of Risk and Preference-Based Values of Safety," *Journal of Risk and Uncertainty* 25(3), 211–232.

Covey, Judith. (2001). "People's Preferences for Safety Control: Why Does Baseline Risk Matter?" *Risk Analysis* 21(2), 331–340.

Jones-Lee, Michael W. (1976). "The Value of Life: An Economic Analysis," London: Martin Robertson and Chicago: Chicago University Press.

Jones-Lee, Michael and Graham Loomes. (1995). "Scale and Context Effects in the Valuation of Transport Safety," *Journal of Risk and Uncertainty* 11(3), 183–203.

McDaniels, Timothy L., Mark S. Kamlet, and Gregory W. Fischer. (1992). "Risk Perception and the Value of Safety," *Risk Analysis* 12, 495–503.

Mendeloff, John and Robert M. Kaplan. (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 Dress.

Savage, Ian. (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.

Slovic, Paul B., Baruch Fischhoff, and Sarah Lichtenstein. (1981). "Perceived Risk: Psychological Factors and Social Implications," in Warner, F. (ed.), *The Assessment and Perception of Risk*. Proceedings of the Royal Society, Val 376, London: The Royal Society, pp. 17–34.

Snedecor, G.W. and W.G. Cochran. (1989). *Statistical Methods, 8th edition*. Ames, IA: Iowa State University Press.

Sunstein, Cass R. (1997). "Bad Deaths," *Journal of Risk and Uncertainty* 14(3), 259–282.

Thomas, K. (1981). "Comparative Risk Perceptions: How the Public Perceives the Risk and Benefits of Energy Systems," in Warner, F. (ed.), *The Assessment and Perception of Risk*. Proceedings of the Royal Society, Val 376, London: The Royal Society, pp. 35–50.

Tolley, George, Donald Kenkel, and Robert Fabian. (1995). *Valuing Health For Policy.* Chicago: University of Chicago Press.

Viscusi, W. Kip. (1979). *Employment hazards: An Investigation of Market Performance*. Cambridge: Harvard University Press.

Viscusi, W. Kip. (1989). "Prospective Reference Theory: Toward an Explanation of the Paradoxes," *Journal of Risk and Uncertainty* 2(3), 235–264.

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber. (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.