BRE
Garston, Watford, WD2 7JR
2000

A risk assessment procedure for health and safety in buildings

G J Raw*, S M Cayless*, J Riley*,
S J Cox† and A Cheyne†

* *BRE Centre for Safety, Health and Environment*

† *Loughborough University*

BRE
constructing the future

DETR
Environment
Transport
Regions

Prices for all available
BRE publications can be
obtained from:
CRC Ltd
151 Rosebery Avenue
London EC1R 4GB
Tel: 020 7505 6622
Fax: 020 7505 6606
email:
crc@construct.emap.co.uk

BR 402
ISBN 1 86081 420 4

First published 2000

Printed from supplied
camera-ready copy

Published by
Construction Research
Communications Ltd
by permission of
Building Research
Establishment Ltd

CONTENTS

A RISK ASSESSMENT PROCEDURE FOR HEALTH AND SAFETY IN BUILDINGS

Gary J Raw[1], Sue J Cox[2], Alistair Cheyne[2], Sandy M Cayless[1] and Jo Riley[1],

[1]BRE [2]Loughborough University

This report describes the development of a procedure for the comparative evaluation of the health and safety risks associated with various aspects of buildings. The purpose of the work was to provide a basis for identifying priorities for action through the Building Regulations, the Housing Fitness Standard and its intended replacement, the new Health and Safety Rating System. The first part of the report describes the derivation of a scale for assigning a severity rating to specific harms that may befall building users. The top four bands of the scale have been adopted for comparing harms that occur, and hazards that exist, within buildings. Data on the frequency of occurrence of the harms were then derived from BRE reports. A procedure was developed for converting the occurrence and severity data, together with the strength of the evidence relating them to the hazard, into a single 'risk index' for each hazard, reflecting known current risks. This was used to place hazards encountered in buildings in a rank order based on this index. Further development of the method indicates that it is sound but that some of the parameters relating to the relative weighting of severity of harm and strength of evidence should be changed.

EXECUTIVE SUMMARY

The purpose of this report is to compare the health and safety risks associated with various hazards in buildings, as a basis for future development of standards and regulations. The report has been compiled as part of the continuing review of technical requirements of the Building Regulations, the Housing Fitness Standard and its intended replacement, the new Health and Safety Rating System.

The report does not cover assessment of what actions would be feasible, or of who should be responsible for carrying them out. The report addresses the issue of risks, which is only one of the factors that should be taken into account when considering the costs and benefits of any proposed actions, legislative or otherwise.

The terms used in risk assessment vary considerably in the literature. In this report, a *harm* is defined as an adverse effect on a person. It might be, for example, a serious illness or injury, but effects on well-being are also taken into account. A *hazard* is a potential cause of harm to a person, for example a faulty staircase. A *hazardous situation* exists when a person is exposed to a hazard, for example by using a faulty staircase. The *risk* associated with a hazard is a function of (a) the likelihood of the hazard causing harm and (b) the severity of the harms or their consequences.

To permit a consistent comparison of health and safety hazards in buildings, a risk evaluation procedure has been developed. The procedure is based on risk matrices, which have been generated for each of the hazards addressed in two published reports, *Building regulation and health* and *Building regulation and safety,* and the publication draft of *Building regulation and fire safety*. The matrices describe the seriousness of the harms to the individual person and the total number of people affected, plus the strength of evidence for the risk. The number of people affected would in some cases be the number of specific occurrences (e.g. number of deaths) and in other cases the number of people affected (e.g. annoyed by noise) on a more continuous basis through the year.

Four classes of harm are used in the matrices; these were derived from the ratings given to stated harms by experts and practitioners in medicine, health and safety. The report describes

two independent approaches that were used to derive the classes of harm. There was good agreement between the results obtained using these two approaches, and with results obtained with non-expert subjects. The resulting four classes of harm are as follows.

- Class I harms are death and other ‘extremely severe' outcomes (e.g. permanent paralysis below the neck, malignant lung cancer, regular severe pneumonia and permanent loss of consciousness).
- Class II is applied to ‘severe' outcomes, for example severe chronic confusion or dementia, mild strokes, regular severe fever and loss of consciousness for hours or days.
- Class III (‘moderate/severe') outcomes include chronic severe stress, mild heart attack, regular severe dermatitis, malignant but treatable skin cancer and regular severe migraine.
- Class IV (‘moderate') includes occasional severe discomfort, chronic/regular moderate skin irritation, benign tumours, occasional mild pneumonia and regular serious coughs or colds.

These outcomes are only examples; a more comprehensive list of harms under each class was available for comparison with the outcomes identified in the reviews. They represent harms as described to the subjects in the studies and do not always correspond to the exact medical terms that might be used.

For each hazard, a ‘risk index' has been calculated. The index is based on the whole risk matrix, but weighted such that Class I harms have by far the greatest influence (the weightings given to Classes I to IV are 100,000, 2500, 50 and 1 respectively). A greater weighting is also given where the evidence is stronger (5 for high, 3 for medium and 1 for low strength of evidence). The indices are then based on the logarithm of the ‘seriousness x number affected' values; this renders the process less sensitive to the exact values given to the weightings.

From the risk indices, it is possible to draw up a rank-ordered list of issues to show which ones represent the highest risk and which would, on this basis, have some priority for new action (other than research). An initial list is presented in Table 1. It is important to bear in mind that the risks associated with a particular hazard may be low because current controls are working successfully: this would not be an argument for relaxing controls. Alternatively a low rating might reflect a lack of evidence, and would therefore not indicate a cessation of research. There is also considerable variation among people in their susceptibility to the effects of each hazard; even hazards with a low overall risk may be significant for some individuals.

Further work to refine the method is described in the report. This work shows that the overall method is sound, and that the hazard ratings are well correlated with hazard ratings given by experts and members of the public who have access to the risk matrices without being told the risk assessment procedure. Also, subjects on the whole did not greatly alter weightings when applying them to different age groups, suggesting that they did not value lives differently for different age groups. However, certain parameters of the method should be altered:

- weightings should be of 1, 10, 30 and 1000 for Classes IV (least severe) to I (most severe) respectively;
- weightings of 1, 2 and 4 be adopted for low, medium and high strengths of evidence.

These new weightings alter the hazard ranks, mainly in the middle range of risk (levels 2 and 3 in Table 1).

The risk index has proved useful in generating a rank order of hazards, but its limitations have to be recognised. It does not at this stage represent a formal risk assessment and it requires further development and a supply of better quality data. It is nevertheless a basis for comparison of risks.

Table 1. Health and safety hazards, grouped in rank order of risk

HOMES	NON-DOMESTIC BUILDINGS
Highest risk	
Hygrothermal conditions Radon[1] Slips, trips & falls on stairs & ramps Slips, trips & falls on the level Burn/scald hazards (other than in house fires) Fire hazards House dust mites Environmental tobacco smoke Slips, trips & falls from buildings Drowning hazards Carbon monoxide Slips, trips & falls involving baths Fire hazards	Slips, trips & falls on the level Radon[1]
Second level of risk	
Electrical hazards Fungal growth Security and the effects of crime Architectural glass Noise Collision/entrapment involving windows Explosions Lead[1]	Slips, trips & falls on stairs & ramps Environmental tobacco smoke Slips, trips & falls on from buildings Sources of infection other than sanitary accommodation Electrical hazards Explosions Carbon monoxide
Third level of risk	
Sanitary accommodation[2] Collision/entrapment involving doors Sources of infection other than sanitary accommodation[2] Space Volatile organic compounds[3] Falling objects Oxides of nitrogen[3] Particulates[4]	House dust mites Volatile organic compounds Lighting Space Hygrothermal conditions Falling objects Collision/entrapment involving doors Fungal growth
Fourth level of risk	
Sulphur dioxide and smoke Landfill gas Pesticides[2]	Sanitary accommodation[5] Noise Collision/entrapment involving windows Collision/entrapment involving lifts Oxides of nitrogen
No clear basis for risk assessment	
Lighting, Electromagnetic fields	Burn/scald hazards, Drowning hazards, Slips, trips & Falls involving baths, Security & the effects of crime, Architectural glass, Sulphur dioxide and smoke, Lead, Landfill gas, Particulates[4], Pesticides, Electromagnetic fields

NOTES

1. Geographically localised.
2. Important potential risk but largely controlled by current standards.
3. Low ranking may be due to insufficient research.
4. Recent research suggests this may get a higher ranking in future.
5. Most significant potential risk in hospitals & schools but generally controlled by current standards.

1. INTRODUCTION

1.1 Background

There is a need for those involved with the regulation of the design and construction of buildings, and with health and safety in relation to building use, to be able to identify the possible harmful outcomes of such use and the associated risks. This requires a systematic method of evaluating and profiling the risks associated with specific building hazards. Health and safety hazards can have outcomes ranging from effects on well-being, which are often transient (e.g. feeling uncomfortably cold), through illnesses requiring varying degrees of medical attention to serious disablement and death. The probability of a particular outcome resulting from a particular hazard also varies widely.

In order to make a consistent comparison between building issues, it is useful to have a broad categorisation of the issues, taking into account both the seriousness of the possible health and safety outcomes and their probability. This report outlines the development of such a system in relation to health and safety in buildings. The primary purpose of the work was to provide a basis for identifying priorities for action in relation to new buildings through the Building Regulations and through the control of standards in existing dwellings. The work carried out so far is an initial step and provides a relatively simple procedure in a complex field.

This chapter introduces the key concepts in the work. The next chapter describes the derivation of a scale for assigning a severity to specific harms that may befall building users. Description of the strength of the available evidence is then introduced. Data on the frequency of occurrence of the harms were derived from two published BRE reports, *Building regulation and health* and *Building regulation and safety,* and the publication draft of *Building regulation and fire safety.* A procedure is then described for representing a hazard by a single index of the known current risks, using the occurrence and severity data, together with an assessment of the strength of the evidence relating the two. This procedure is used, in Chapter 4, to place hazards encountered in buildings in a rank order based on the risk currently associated with each hazard. Chapter 5 reports work to refine the methodology and Chapter 6 draws the main conclusions.

1.2 Hazard and harm

This report uses terms related to risk assessment. Such terms are used with different meanings by different authors, and they are therefore defined here as used in this report.

- A *harm* is the adverse effect on a person, which may arise from exposure to a hazard; it is the damage done to individual structure or function. For example, people can suffer harm through physical injury to different parts of their body, or through physical or psychological illness of different types. A harm can range from trivial, such as brief discomfort, to severe illness or injury, and death. Harm is often confused with related concepts such as those of hazard and risk.
- The *consequences of a harm* define its magnitude or severity and represent the reasons why, or the ways in which, it impairs the person's quality of life and normal behaviour in society. A broken leg, for example is a harm which can have the following consequences: it causes pain and discomfort which impair the quality of normal life, it compromises mobility, it forces greater dependency on others, it precludes working, it can cause financial embarrassment, and it interferes with short and medium term goals.
- A *hazard* is a potential cause of harm to a person, for example low ambient temperature or a faulty staircase.
- A *hazardous situation* exists when a person is exposed to a hazard, for example by using a faulty staircase.

The nature and severity of harms to individual building 'users', and the consequences of these harms, are highly varied. The extent of any harm is dependent on the nature of the hazardous event and the vulnerability of the person involved in the event. The consequences of the harm are dependent on the circumstances of the individual. For example, the consequences of a disabling injury for a single parent who has young children to support may be different in nature and degree to the consequences for a person who has no dependants.

Logically, there will be several different dimensions of consequence which can be used to describe harm and measure its severity. Some of these dimensions may prove to be common to many different types of harm, and thus offer a series of scales on which harms can be measured and compared. The concept of harm is also affected by the context within which it occurs: in relation to health and safety it is framed by the nature of the hazardous situation.

1.3 Risk

The relationship between hazard (or hazardous situation) and harm is often referred to in terms of 'risk'. Risk has been defined as a function determined by two variables: (a) the likelihood of a particular hazard causing harms and (b) the magnitude (severity) of the harms or their consequences (e.g. Cox & Tait 1991). There are several derivations of this function which can be logically or empirically justified (e.g. Cox & Cox 1992). The question is, how can risk - so defined - best be described in relation to building hazards?

It has been argued (Cox & O'Sullivan 1995) that any situation involving possible harm to building users may logically involve more than one type of hazard and more than one type of harm. Thus there may be different probabilities of occurrence associated with different severities or magnitudes of each type of harm or consequence. For practical purposes, if a limited number of the most probable types of harm can be identified, then particular building hazards could be assessed in terms of risk of each of these harms occurring.

Harms can be described using a two dimensional 'risk matrix' of probability by severity, of the form suggested by Cox & Tait (1991) and further described by Cox & O'Sullivan (1995). The risk matrix can provide a simple visual representation of risk by plotting the probability of occurrence of different severities of harm for any particular hazard. This has been done in the reports *Building regulation and health* and *Building regulation and safety* and the matrices are reproduced in Appendix J.

The overall probability of a particular harm occurring is the product of a number of other probabilities which together describe the chain of events which link hazard to harm. For example the probability of a person breaking a leg as a result of a fall while using stairs in a building is the product of:

- the probability of a hazard (stairs) being present in a building;
- the probability of a hazardous situation arising when the hazard is present (a building user climbing the stairs);
- the probability of an accident occurring (e.g. a person slips and falls whilst on the stairs);
- the probability of a particular harm occurring (a person breaks a leg as a result of the fall).

These probabilities are obviously conditional. The probability of a hazard occurring in a building is, in many cases, high (e.g. doors, stairs, WCs, airborne fungal spores), and so is the probability of a building user interacting with any or all of them. In other cases the probability is lower (e.g. temperatures below 12°C, radon above the action level). The probabilities of the various hazardous events associated with these hazards occurring varies still more, as do the probabilities of particular types of harm arising from these accidents.

It is not always possible to make reliable enough estimates of the separate probabilities listed above, and therefore the likelihood dimension of the matrix is often approximated in different ways. This report is concerned with the overall risk in the population since the separate probabilities are not sufficiently well specified. The expressions of risk therefore relate to the likelihood of harm from a specified hazard in a specified population. For example, approximately 4,500 people per year in England and Wales die as a result of a home accident. If the total population in England and Wales is taken as 48 million, we can calculate the risk of death due to an accident in the home to be approximately 1 in 10,000 per annum. In this report, the total population is normally treated as a constant; the risk is therefore defined by the number of people affected.

1.4 Overview of report content

This report describes initial attempts to address the following:

- identification of harms that may occur in buildings;
- rating the severity of those harms;
- derivation of the likelihood, and the strength of the evidence, that each of a range of hazards will result in harms at each level of severity;
- assigning numerical weights to severity of harm and strength of evidence;
- calculating a risk index which permits hazards to be placed in rank order of risk to health and safety.

2. GENERATION OF A RATING SCALE FOR SEVERITY OF HARM

2.1 Introduction

This chapter summarises a study concerned with measuring the severity of harms which may befall building users. The aim was to develop a scale for measuring severity of harm which may be used in the evaluation of health and safety risks in buildings in relation to the Building Regulations, the Housing Fitness Standard and its intended replacement, the new Health and Safety Rating System.

Existing measures of harm deal only with specific examples or groups of injuries and/or diseases, while health profiling methods measure the effects of disease and injury, without attempting to quantify harms. No existing scale deals specifically with the full range of harms likely to be sustained in buildings.

The first stage of the work was to produce a list of harms which are representative of those which might be sustained in buildings; the method of generating the list of harms is described in Appendix A and the harms are listed in Appendix B.

Two approaches to the development of a measure of severity of harm were then explored; the remainder of this chapter is a summary of this study, which is described in detail in Appendix C. The first approach involved the direct comparison and rating of the severity of harms to construct an ordinal scale. The second approach involved identification of the consequences of harm and then rating types of harm against scales developed from those consequences.

2.2 First approach: direct ratings of severity

2.2.1 Method

A pilot study tested and refined the procedures to be used in the main study. Subjects rated representative harms, thus allowing the researchers to explore the criteria used and to collect preliminary data on the distribution of harms. The main rating exercise was then carried out and each harm assigned a mean severity rating. Finally, a reliability test was carried out.

Subjects were drawn from a range of health and safety and medical fields. A list of 293 harms (see Appendix B) was used in the pilot, reduced to 255 in the main study by combining harms with similar meanings and mean ratings, and removing some harms with the greatest agreement among the subjects. Each harm was rated against a seven point scale. The following scale was used in the main study: [1] Extremely Slight, [2] Slight, [3] Moderate/Slight, [4] Moderate, [5] Moderate/Severe, [6] Severe, and [7] Extremely Severe.

In the pilot study, all of the 12 subjects confirmed that they assessed each harm against its most likely consequences/outcomes (e.g. how long it would take to recover, if recovery were possible) when rating its severity, rather than just making a quick global decision. This approach was explicitly requested in the instructions for the main study.

2.2.2 Results

Each harm is shown in ascending order of mean severity rating in Appendix D, based on the combined responses from all 32 subjects taking part in the pilot study or the main study. The level of agreement between subjects was high for many items, but quite low for others; this may have been because of variation in the amount of information in the descriptions of the harms.

Table 2 shows seven different broad outcome categories and the harms relating to them. This table has been constructed using those items which clearly mark either one end of the scale or the centre of one of the original bands (i.e. 1.5, 2.5, etc), and which showed a low variance of

ratings. Separating out categories of harm, as in Table 2, presents scales that are more easily appreciated. However, there are no example harms in some cells and consequently each scale is incomplete and could not, therefore, be used on its own.

2.2.3 Reliability test

Seven harms were used to test the reliability of the scale using a questionnaire format. Subjects were asked to rate each of the scale markers on two occasions, separated by at least 24 hours.

No indication was given to subjects on the first occasion that they would be asked to complete it once more. The rank order of all the harms remained the same and all reliability coefficients are of an acceptable level. The highest coefficients are at the two extremes of the scale.

2.2.4 Conclusions on direct ratings of severity

The first approach to the development of a measure of severity of harm provides a simple scale. This method assumes however that the concept of harm is unidimensional and that representative values of severity of harm can be identified and attributed with confidence. This approach can be defended if different consequences of harm can be identified and statistically combined to form a single factor, and if the resulting scale correlates highly with that found in the first approach. The following section describes the development of such a scale.

2.3 Second approach: scaling by consequence

2.3.1 Introduction

The second approach to the development of severity ratings was carried out in two phases. First, consequence constructs were identified using a method known as ‘triadic elicitation'. The harms were then rated using consequence scales developed from these constructs. The results of this exercise were analysed and a scale based on consequence ratings produced.

2.3.2 Generation of constructs

The list of harms was stripped of all modifying adjectives (e.g. ‘slight', ‘regular') to produce a list of 126 basic harms. Each of 30 subjects evaluated 21 sets of three harms (triads). The subjects said why two out of each set were similar to each other, but differed from the third, in terms of their consequences. The description of similarity and difference represents a ‘construct'. Five basic constructs were thus derived: Results in Death, Changes Lifestyle, Restricts Mobility Permanently, Affects Physical Performance (in or out of work) and Inhibits Social Interaction. These consequences resemble some of the outcomes identified in previous research, as well as the dimensions involved in health profiling.

2.3.3 Rating harms by consequences

One harm was taken from each scale point and each mid-point, representing, as much as possible, the broad outcome categories set out in Table 2. The 13 resulting harms were each rated on each of the five consequence constructs. The data were analysed using principal components analysis, which produced one primary factor, accounting for 76.9% of the overall variance. The fact that one primary factor could be derived endorses the unidimensional methodology used in the first approach.

Adding the scores from each consequence scale provides a new score ranging from 5, the absence of negative consequences, to 35, the presence of all negative consequences. An alternative scale was derived by asking subjects to rate how important each consequence construct was when judging severity of personal harm. Multiplying each construct score by its importance weighting and then adding these weighted scores gives an overall severity score for

each harm, which was standardised to create a seven point scale. This scale was then compared to one derived from the simple addition of consequence scales. A very high correlation between the two scales indicated that there was no real difference between them.

2.4 Scale comparison

Comparing scale values from the two approaches showed that all outcome scale values are within one scale point of their corresponding card sort values. The markers at the ends of the scales match very closely. The scales used were, however, not identical and relative ranks may provide a better basis for comparison. Those items in the lower and mid parts of the scale changed their order slightly but there is a high level of agreement between the scales. Therefore the hypothesis that there is a single underlying factor of severity of harm is supported, and a single rating scale can be used.

2.5 Comparison of expert and public ratings

A further study was conducted, in which both specialists in the field of health and safety ("experts") and the general population ("the public") rated a reduced list of representative harms (see Appendix E). There was a high degree of agreement between experts and non-experts, and between the experts and the expert group in the earlier study, who had rated a larger number of harms. This contrasts with earlier studies of perceived risk, which generally do show a difference between experts and the public. Thus, it seems likely that differences in perceived risk between experts and the public are due to differences in estimated likelihood or differences in the function used to combine severity and likelihood.

The similarity between experts and the public would seem to suggest that the scale derived is relatively robust and forms a sound basis for the comparative evaluation of health and safety risks in buildings.

2.6 Conclusions on generation of a rating scale

The ranges covered by the 7 scale points (1-1.49, 1.5-2.49, 2.5-3.49, 3.5-4.49, 4.5-5.49, 5.5-6.49 and 6.5-7) can be used as 7 bands of severity of harm. The lower two bands would have relatively little value for large scale assessment of risks, being composed of harms for which statistics would not be kept, and which are at a level unlikely to affect legislation. These levels of harm would however have significance to individuals who would normally seek to avoid them, and therefore would benefit from guidance from Government or other sources. The next band (2.5-3.49) would be of variable value, depending on the specific group of harms being considered.

The recommendation is therefore that the top four bands of the scale be adopted for the purpose of comparing harms that occur, and hazards that exist, within buildings. A five-band scale could be used for more local evaluations, for example involving specific buildings or regions, or in assessments related to well-being and productivity. The five-band scale is shown in Table 3, with examples of the harms that fall in each band.

Table 2. Comparison of rated harms from different outcome categories

SCALE POINT[1]	PSYCHOLOGICAL/ BEHAVIOURAL	MUSCULO-SKELETAL	CARDIOVASCULAR	SKIN-RELATED	ONCOLOGICAL	RESPIRATORY	NEUROLOGICAL
1				Slight graze to hand			
1-2	Occasional slight discomfort						
2-3	Occasional slight restriction of social behaviour	Temporary slight backache				Occasional mild cough/cold	
3-4	Temporary immediate moderate stress		Short term slight high blood pressure			Occasional mild pneumonia	Occasional moderate migraine
4-5	Chronic delayed moderate stress		Mild angina	Moderate burns/scalds to torso	Benign tumours	Chronic/regular mild infection of the lung	Immediate moderate tinnitus
5-6	Constant serious anxiety		Immediate mild heart attack		Malignant skin cancer - treatable	Regular severe fever[3]	
6-7		Loss of arms		80% burns and scalds	Malignant lung cancer	Regular severe pneumonia	
7	Suicide	Permanent paralysis below neck[2]	Fatal stroke/heart attack		Leukaemia - non-treatable		Permanent loss of consciousness

[1][1] Extremely Slight, [2] Slight, [3] Moderate/Slight, [4] Moderate, [5] Moderate/Severe, [6] Severe, and [7] Extremely Severe

[2]May alternatively be classified as neurological.

[3]Not specific to respiratory disease.

Table 3. Comparison of rated harms from different outcome categories (five classes)

CLASS	SCALE POINTS	PSYCHOLOGICAL/ BEHAVIORAL	MUSCULO-SKELETAL	CARDIOVASCULAR	SKIN-RELATED	ONCOLOGICAL	RESPIRATORY	NEUROLOGICAL
V Minor	2.00 to 3.49	Occasional slight restriction of social behaviour	Temporary slight backache	Short term slight high blood pressure	Occasional slight dermatitis		Occasional mild cough/cold	Delayed slight tinnitus
IV Moderate	3.50 to 4.49	Occasional severe discomfort	Moderate ligament injury in hip		Chronic/regular moderate skin irritation 10% burns/scalds	Benign tumours	Occasional mild pneumonia Regular serious cough/cold	Brief loss of consciousness
III Moderate/ Severe	4.50 to 5.49	Chronic immediate severe stress	Temporary paralysis of arms[1]	Immediate mild heart attack	Severe burns/scalds to hands Regular severe dermatitis	Malignant skin cancer - treatable	Occasional moderate pneumonia	Immediate severe tinnitus Regular severe migraine
II Severe	5.50 to 6.49	Severe chronic confusion/dementia	Chronic severe backache Loss of foot	Mild stroke			Regular severe fever[2]	Loss of consciousness for hours/days
I Extremely Severe	6.50 to 7.00	Suicide	Permanent paralysis below neck[1]	Fatal stroke/heart attack	80% burns and scalds	Malignant lung cancer – non-treatable	Regular severe pneumonia	Permanent loss of consciousness

[1]May alternatively be classified as neurological.
[2]Not specific to respiratory disease.

3. ASSESSING THE STRENGTH OF THE EVIDENCE

3.1 Introduction

One key objective in the measurement of risk is establishing the relationship between hazardous situations (causes) and the harms (effects). It is therefore not sufficient to estimate the number of people affected by harms of specified severity. Some weight must be attached to the strength of the evidence that the estimated numbers are affected (or indeed whether any are affected at all). This is represented in the risk matrices by having one to three stars in the relevant matrix cells: the strength of the evidence for a particular hazard-outcome relationship is classified as follows:

- *high*: a causal effect is well established, affecting the number of people indicated (shown as ******* in the matrix for each issue);
- *medium*: probable causal effect but less well-established, e.g. causal effect at extreme exposures and indirect evidence of effect at lower exposures (****** in the matrices);
- *low*: possible causal effect, e.g. a plausible link but with limited supporting evidence (* in the matrices).

The remainder of this chapter describes the conceptual basis for assigning the stars and for using the evidence in regulatory decisions. It is derived to a large extent from the first edition of *Building regulation and health* (1986).

3.2 The scientific status of different types of evidence

The type of evidence available on building and health issues is highly variable. Ideally, in order to demonstrate cause and effect, the following findings would have been made in reliable studies:

- a correlation between the prior presence of the hazard and the occurrence of the health outcome, if possible with a dose-response relationship (i.e. the higher the level of the hazard, the greater the probability or seriousness of the health outcome);
- well-documented case studies in which the specified hazard was the only plausible cause of the health outcome;
- a plausible or proven causal mechanism;
- experimental evidence that removing the hazard reduces the health problem;
- experimental evidence that (accidentally or unwittingly) introducing the hazard causes the health problem.

The more specific the relationship the more useful the evidence will be. Such evidence should ideally be based on both laboratory studies and field trials or epidemiology (particularly longitudinal epidemiology). Studies of animals contribute, particularly to establishing physiological mechanisms, but most of the work would have to be on illness and injury to people.

Correlational evidence is often the easiest to collect, but does show only correlation, not cause (e.g. sunburn and eating ice cream are correlated but ice cream does not cause sunburn). However, in the absence of a correlation, it is difficult to be convincing about cause. Data on a range of possible hazards and modifying factors (e.g. age, sex) can be taken into account in calculating the correlation with the each hazard. This use of multivariate statistics is important, if not essential in establishing the credibility of the correlation.

The evidence is strengthened if there is a dose-response relationship. For example in the case of radon the evidence is derived partly from an occupational setting (mainly from studies of uranium miners) and a relationship has been established between radiation dose and the probability of acquiring lung cancer.

Correlational evidence is further strengthened by the existence of a plausible or even proven causal mechanism. For example there are good physiological explanations of why low temperature should increase death rates from the specific causes observed in excess in winter. The existence of a ‘plausible causal explanation' is very important in this sort of situation where experimental trials are difficult to perform and limited to studying the impact of housing improvements on health.

The use of experimental interventions is the most stringent single test of cause and effect, but is difficult to do, particularly in ‘real life' situations as distinct from short-term laboratory trials. Nevertheless, in a few cases such evidence is available. Less rigorous evidence can be obtained where there is a ‘natural experiment' resulting from a change in practice (e.g. increasing installation of central heating) but it can be very difficult to take into account other factors which change over the same time period.

In the case of diseases that are rare or newly discovered, the only available evidence may be case studies. Although these have an important role in establishing the nature of the health outcome, and identifying likely causes, they are generally of limited use for drawing conclusions about the implication of particular building-related health issues in the population as a whole.

In the case of many hazards, particularly safety hazards, the link between cause and effect is, in general terms at least, not in doubt. The primary area of uncertainty is then the number of people affected. The data derived from accident surveillance systems are therefore critical in the estimation of risk. Such systems show that accidents affect a large number of people. However, due to the nature of various methods of data collection, these numbers may well under-represent the number of deaths caused by accidents. For example, one important source of information is the death certificate but these certificates do not always record the cause of death in a way that clearly identifies any building-related element in the cause.

Much of the data available in relation to health and safety in building have not been collected for investigations of cause and effect (much of these data are collected for the purpose of establishing culpability). They therefore provide a weak basis for such investigations, and conclusions drawn from these data can be, at the best, only tentative. What is required is a series of dedicated studies on the cause and effects of particular building illnesses and accidents. It is obvious from the available data that behavioural as well as structural factors need to be studied in relation to causes, and both long- and short-term consequences in relation to effects.

Establishing the role of a specific housing factor such as overcrowding, in the face of many other possible contributing factors, is difficult even when the associated disease or injury is well defined. In many situations (e.g. the investigation of noise as a causal factor in ill health) neither the possible disease nor the probable exposure is fully defined, and any effect is likely to be exerted in combination with other contributory factors. In this case it is important to recognise that direct evidence is unlikely to be produced, no matter how strong the effect.

Evidence from scientific research is invaluable in reaching rational decisions on the prevention of ill health. However, it is important not to confuse importance with ease of measurement, nor to assume that lack of evidence indicates the absence of a problem: it may be that the data are not available because the necessary research is too difficult or too costly to undertake.

3.3 Applying evidence to regulatory decisions

For the purpose of justifying a regulation or standard on health grounds it is necessary to establish that the regulation or standard is intended to avert some outcome which is:

- correctly identified as a health problem;

- sufficiently important to merit public action (this is a function of the seriousness of the health outcome and the proportion of the population affected);
- likely in practice to be averted by the legislation.

The critical statistic is the improvement in the health of the population that is predicted to result from the change in building or operating practice. Unfortunately there are few examples where the full calculation is possible, since information would be required on:

- the probability of exposure to a hazard (e.g. probability of a failure in a heater resulting in carbon monoxide gas reaching 10 ppm;
- the probability of specific health outcomes (e.g. coma, pneumonia) occurring as a result of exposure to the hazard;
- the confidence which can be placed in the above statements of probability, based on the strength of the evidence;
- the seriousness of the health outcomes which are identified;
- the expected success rate of the building standard or regulation in eliminating the hazard.

Usually at least part of the information is missing or can only be guessed. Therefore risks to health should be considered even if proof of the degree of risk cannot be obtained, and legislation must sometimes proceed even with incomplete knowledge.

To complicate matters yet further, it is sometimes necessary to have separate data for different groups of buildings or people. For example in making decisions about Building Regulations (which apply to new buildings), information about hazards associated with newer homes would often be more useful than information about the building stock as a whole. Alternatively, if a particular hazard were considered to present a measurable risk to only one group in the population (e.g. the elderly) then information on this group and the homes that they occupy would be critical.

The variation in risk between different groups in the population is a major problem in formulating appropriate action. Although a ‘high risk' group can often be identified, it is unusual for this group to contribute more than a small proportion of the total population risk. The real value of quantitative risk estimates is the opportunity to aim for consistency in setting standards for different hazards. Public health should include health of susceptible and vulnerable persons, i.e. old, young, ill, socially disadvantaged as well as the ‘average' or ‘normal' person;

Thus, decisions must sometimes be taken on the grounds of plausibility of a causal link and on the basis of the indirect evidence. For example the decision about the extent to which party-wall insulation would reduce illness must be based on the surveys of annoyance and on the plausibility of the link between noise, disturbed sleep and illness rather than any direct observation of cause and effect.

4. FIRST RISK EVALUATION

4.1 The procedure

This section describes a risk evaluation carried out using data on severity of harm and strength of evidence, as presented in the preceding chapters. Subsequently, some refinements have been made to the procedure, as described in Chapter 5. This chapter records the original risk evaluation as it was performed, using the data available at the time. Data were derived as described in two published reports, *Building regulation and health* and *Building regulation and safety*, and the working draft of *Building regulation and fire safety*. These reports summarise the risk data in matrices (see Table 4 and Appendix J).

Table 4. Example of matrix summary of building-related health and safety risks

	NUMBER OF PEOPLE AFFECTED IN THE UK PER YEAR					
OUTCOME	100,000+	10,000+	1000+	100+	10+	1+
Class I					**	***
Class II				***		
Class III		*	**			
Class IV	***					

Outcome classes I to IV correspond to different severities of harm, as set out in Table 3. No attempt has been made to assign different weights to different lives. Thus, the death of an elderly person with a residual life expectancy of perhaps five years is assigned the same weight as the death of an infant.

The risk evaluation procedure combined data on the severity of the harms resulting from a particular hazard, the number of people affected and the strength of the evidence (number of stars) relating the hazards to the harms. A single rating for each hazard was derived from each risk matrix presented in the companion reports, as follows.

1. For each line (each class of harm), select the leftmost cell which has at least one star and assign a value of 1 for one star, 3 for two stars and 5 for three stars.
2. Multiply this value by the number of people affected (between 1 and 100,000). The orders of magnitude from the risk matrices have been used except for domestic safety issues; in this latter case exact figures have also been used since the available data provide a more precise figure than for other issues.
3. Multiply the resulting figure by a value representing the seriousness of the harm (Extremely Severe = 100,000, Severe = 2500, Moderate/Severe = 50, Moderate = 1).
4. Sum the products for the four harm classes and take the natural logarithm of the sum.
5. Divide by the maximum possible score (i.e. the score that would be obtained with three stars in the “100,000+” column in each row).
6. Multiply by 100 to obtain a figure analogous to a percentage.

The values generated by this procedure are referred to here as a ‘risk index'. They not formal estimates of risk but a higher index should represent a higher risk. Although based, where possible, on actual numbers of harms, the index is most useful for placing hazards in rank order, rather than stating the actual risks attributable to a particular hazard. The indices should therefore not be assigned an absolute value; they are useful only for placing the issues in a rank order.

4.2 Results

The procedure described above resulted in the health and safety hazards being placed in rank order as shown in Tables 5 and 6. The risk indices relate to the population as a whole. Individuals will exist who are either more likely to be exposed to the hazard, or more susceptible to its effects. A low risk index should not be taken to mean that no risks exist, only that they are either small or that there is insufficient evidence at present that there is anything other than a small effect. Neither should the indices be taken to be an indicator of what action is necessary.

Table 5. Rank order of risks for health and safety hazards (homes)

HAZARD	INDEX	RANK	EXACT INDEX[1]
Hygrothermal conditions[1]	84	1	-
Radon	81	2	-
Slips, trips and falls on stairs and ramps	79	3	80
Slips, trips and falls on the level	73	4	77
Burn/scald hazards	75	5	76
Fuel gas[2]	75	6	76
House dust mites	75	7=	-
Environmental tobacco smoke	75	7=	-
Slips, trips and falls from buildings	70	9	72
Drowning hazards	70	10	71
Slips, trips and falls involving baths	64	11	70
Carbon monoxide	70	12	-
Electrical hazards	67	13	67
Fungal growth	66	14=	-
Security & the effects of crime	66	14=	-
Architectural glass	64	16	66
Noise	64	17	-
Collision/entrapment involving windows	58	18	64
Explosions	63	19	63
Lead	61	20	-
Sanitary accommodation	58	21	-
Collision/entrapment involving doors	54	22	58
Other sources of infection	57	23=	-
Space	57	23=	-
Volatile organic compounds (VOCs)	57	23=	-
Falling objects	53	26	57
Oxides of nitrogen	56	27	-
Particulates	50	28	-
Sulphur dioxide and smoke	49	29	-
Landfill gas	48	30	-
Collision/entrapment involving lifts[3]	41	31	44
Pesticides	26	33	-
Lighting; Electromagnetic fields	-	-	-

[1] For safety hazards, two figures are given. The first is based on order of magnitude number of accidents (and is therefore more comparable with the health hazard indices); the second is based on the best estimate of actual number of accidents. The ranks for safety hazards are based on the higher of the two indices.
[2] Temperature, humidity & air velocity.
[3] Includes carbon monoxide poisoning, fires, burns and explosions involving fuel gas.
[4] Covers both domestic and non-domestic premises.

Table 6. Rank order of risks for health and safety hazards (non-domestic)

HAZARD	INDEX	RANK
Slips, trips and falls on the level	72	1
Radon	70	2
Slips, trips and falls on stairs and ramps	68	3
Environmental tobacco smoke	67	4
Slips, trips and falls from buildings; Other sources of infection	66	5=
Electrical hazards	64	7
Explosions	63	8
Carbon monoxide	61	9
House dust mites	58	10
Volatile organic compounds (VOCs)	57	11
Space; Lighting	54	12=
Hygrothermal conditions[1]; Falling objects	53	14=
Collision/entrapment involving doors	51	16
Fungal growth	50	17
Sanitary accommodation	49	18
Noise	48	19
Collision/entrapment involving windows	46	20
Oxides of nitrogen	37	21
Burn/scald hazards; Drowning hazards; Slips, trips and falls involving baths; Security & the effects of crime; Architectural glass; Lead; Sulphur dioxide and smoke; Landfill gas; Particulates; Pesticides; Electromagnetic fields	-	-

A low risk index should not be taken to mean that no risks exist, only that they are either small or that there is insufficient evidence at present that there is anything other than a small effect. Neither should the indices be taken to be an indicator of what action is necessary.

An issue with a high index may offer little potential for improvement through changes to current building practice, environmental tobacco smoke for example. In other cases, the rank order could be low either because current practice is controlling what would otherwise be a high risk (e.g. sanitary accommodation) or because the required action is research to clarify what the risks are (e.g. VOCs). In still other cases there will be both a high current risk and the potential for action; in these cases, the risk evaluation calls for a more detailed exploration of the possibilities, either for changes to practice or for research into what changes would be most appropriate.

The ranks given will also depend on the grouping of hazards. For example, in Table 5 an index has been shown for fuel gas as a hazard, regardless of whether health, safety or fire safety are considered. This clearly gives a higher index than, for example, carbon monoxide on its own.

Where action is required, it could take a number of forms, related to building design/remediation or activities within buildings. Alternatively, where a low index value is due primarily to lack of data, the action indicated might be further research. It is also important to note that this rank ordering does not take into account possible reasons for action other than health and safety (e.g. energy conservation), cost-effectiveness, or the technical problems that may represent barriers to action.

Since there are fewer deaths associated with non-domestic buildings, the ratings are lower overall; this limits the usefulness of the scheme for comparing issues within the non-domestic sector. In many cases, action in non-domestic buildings will be guided not by legislation related to health and safety, but by the requirement for an environment in which people can work efficiently; this would normally require a higher degree of refinement than that required for health purposes.

4.3 Sensitivity analysis

The values employed in the procedure to represent degrees of harm and the strength of evidence are clearly not fixed and alternative values could be chosen. The effect of the assumptions made was tested using the figures for health in homes, since these figures provide the widest range of indices. The results are shown in Table 7 (the first column shows the standard indices).

Table 7. Sensitivity analysis

HAZARD	RISK INDEX (RANK)*					RANGE OF RANKS
	1	2	3	4	5	
Hygrothermal conditions	84 (1)	83 (1)	85 (1)	83 (1)	86 (1)	0
Radon	81 (2)	79 (2)	83 (2)	82 (2)	81 (2)	0
House dust mites	75 (3=)	73 (3=)	77 (3=)	74 (3)	77 (3)	½
Environmental tobacco smoke	75 (3=)	73 (3=)	77 (3=)	73 (4)	76 (4)	½
Carbon monoxide	70 (5)	67 (5)	72 (5)	70 (5)	70 (5)	0
Fungal growth	66 (6=)	64 (7=)	68 (6)	65 (6=)	67 (6=)	1½
Security	66 (6=)	66 (6)	65 (7)	65 (6=)	67 (6=)	1
Noise	64 (8)	64 (7=)	62 (8)	63 (8)	65 (8)	½
Lead	61 (9)	59 (10)	59 (12)	62 (9)	60 (9)	3
Sanitary accommodation	58 (10)	56 (11=)	60 (9=)	57 (10=)	59 (10=)	2
Other sources of infection	57 (11=)	56 (11=)	60 (9=)	56 (12)	59 (10=)	2
Space	57 (11=)	60 (9)	54 (14=)	57 (10=)	58 (12=)	5½
Volatile organic compounds	57 (11=)	55 (14)	60 (9=)	55 (13=)	58 (12=)	4½
Oxides of nitrogen	56 (14)	56 (11=)	54 (14=)	55 (13=)	57 (14)	2½
Particulates	51 (15)	47 (15=)	55 (13)	50 (15)	52 (15)	2½
Sulphur dioxide and smoke	49 (16)	47 (15=)	52 (16)	48 (16)	50 (16)	½
Landfill gas	48 (17)	45 (17)	51 (17)	47 (17)	49 (17)	0
Pesticides	26 (18)	24 (18)	24 (18)	25 (18)	26 (18)	0

1 = Standard index; 2 = Severity weightings altered to 1, 10, 1000, 10000;
3 = Class I severity weighting altered to 1000000; 4 = Strength of evidence weightings altered to 1, 5, 9;
5 = Strength of evidence weightings altered to 1, 2, 3.

There was little effect on ranking when the values assigned to classes of harm were made more similar (column 2). The only significant change is that issues relating primarily to large numbers of less serious harms move slightly higher up the rank order. Changing the weighting for Class I to 1,000,000 (column 3) also produces relatively small changes in rank, in the opposite direction to column 2. Similarly, changing the values given to the strength of the evidence (columns 4 and 5) had minimal effect on the indices, as shown in columns four and five of Table 7.

We can conclude that the dominant effect on the indices is the number of occurrences of each category of harm which have been attributed to each hazard. The greatest variation is in the middle of the rank order, perhaps reflecting the fact that the greatest and least hazards are the easiest to identify. Uncertainty in the ranking must be considered in using the ranks in any policy decisions and further work is needed to increase confidence in the weightings.

4.4 Linear index

The risk index indicates rank order but does not represent ratios of risk. As stated above, a rank ordering is probably all that can reasonably be expected at this stage of the work and it could be misleading to assign absolute values to the indices. Nevertheless, it may be instructive to

examine the approximate relative magnitudes of risk. This can be done simply by excluding the step of taking logarithms (Step 4 in Section 4.1) and entering a multiplier to obtain a manageable range of values (most between zero and 1000). The values of this 'linear index' (see Table 8) should be read with caution and little should be read into differences that are too small to show up in the logarithmic index. The values do however show that there is a wide range of estimated risks, as a basis for judging the relative importance of health and safety hazards.

Table 8. Linear risk index for health and safety hazards

HAZARD	HOMES[2]		NON-DOMESTIC
	ORDER OF MAGNITUDE	EXACT	
Hygrothermal conditions	2107	-	1
Radon	980	-	59
Slips, trips & falls on stairs and ramps	517	665	35
Slips, trips & falls on the level	122	344	113
Burn/scald hazards	229	289	-
Fuel gas[1]	196	251	-
House dust mites	220	-	3
Environmental tobacco smoke	206	-	26
Slips, trips & falls from buildings	61	92	25
Drowning hazards	69	71	-
Slips, trips & falls involving baths	13	68	-
Carbon monoxide	61	-	6
Electrical hazards	32	33	13
Fungal growth	23	-	0
Security & the effects of crime	21	-	-
Architectural glass	13	20	-
Noise	13	-	0
Collision/entrapment involving windows	3	13	0
Explosions	10	12	10
Lead	6	-	-
Sanitary accommodation	3	-	0
Collision/entrapment involving doors	1	3	1
Other sources of infection	3	-	20
Space	3	-	1
Volatile organic compounds	2	-	2
Falling objects	1	3	1
Oxides of nitrogen	2	-	0
Particulates	1	-	-
Sulphur dioxide and smoke	0	-	-
Landfill gas	0	-	-
Collision/entrapment involving lifts	-	-	0
Pesticides	0	-	-
Lighting	-	-	1
Electromagnetic fields	-	-	-

[1] Includes carbon monoxide poisoning, fires, burns and explosions involving fuel gas.

[2] Alternative values are given for domestic safety hazards, based on order of magnitude and exact numbers.

5. REFINEMENT OF RISK ASSESSMENT METHOD

5.1 Introduction

The first risk evaluation, as described in the previous chapter, was required at a time when a number of aspects of the risk assessment methodology had been only partly tested (for the publication of *Building regulation and health* and *Building regulation and safety*). Hence, a series of further studies has been conducted to verify and/or refine the following aspects of the procedure:

- overall ratings of hazards
- weightings of severity of harm
- weightings of strength of evidence
- ratings of the most severe harms.

These studies used both experts and non-expert subjects and they therefore also contribute to our understanding of why experts and non-experts differ in their assessments of risk.

5.2 Rating of hazards

5.2.1 Introduction

This study aimed to examine how experts and non-experts rate a list of hazards (rather than harms, as in the previous studies), to determine:

- whether the rank order of hazards given by the risk assessment methodology accords with the rank order given independently by non-experts and experts in health and safety fields;
- whether experts and non-experts differ in their ratings; and
- the extent to which effect the rankings given by the two subject groups are affected by the amount of information provided about the hazards.

A more detailed description of the work is given in Appendix F.

5.2.2 Method

The subjects from each group were divided among three conditions:

1: those shown a completed risk matrix (see example in Section 4.1) for each hazard and told to take into account the number of people affected, the severity of the outcome and the strength of the evidence linking the outcome to the hazard;

2: those shown an empty risk matrix with explanatory text and told to take into account the number of people affected, the severity of the outcome and the strength of the evidence;

3: those not shown a risk matrix but advised to take into account the number of people affected, the severity of the outcome and the strength of the evidence.

The study used a reduced list of hazards from Chapter 4, selected at random. Subjects were asked to rate each of the listed hazards against a seven point scale in which 1 equalled “very minor” and 7 equalled “very major”. When all hazards had been rated, subjects were asked to rank the hazards within each of the seven categories. Thus, each subject placed all the hazards in rank order.

5.2.3 Results and discussion

Table 9 shows the rank correlation coefficients for each subject group and condition compared with each other. There is a linear relationship among the three conditions for both experts and non-experts, and between the two subject groups for each condition (Table 10). The correlations between direct ranking and risk assessment rankings increase with the amount of information given and the correlations are higher for experts than for non-experts.

Table 9. Rank correlation coefficients

	Experts			Non-experts		
Condition:	1	2	3	1	2	3
Experts (2)	0.82**	1				
Experts (3)	0.72**	0.76**	1			
Non-experts (1)	0.78**	0.72**	0.52**	1		
Non-experts (2)	0.54**	0.75**	0.61**	0.45*	1	
Non-experts (3)	0.46**	0.72**	0.53**	0.41*	0.91**	1
Risk assessment	0.92**	0.73**	0.66**	0.69**	0.37*	0.30

*p<0.05 **p<0.01

Table 10. Correlation coefficients

	Experts			Non-experts		
Condition:	1	2	3	1	2	3
Experts (2)	0.84**	1				
Experts (3)	0.69**	0.82**	1			
Non-experts (1)	0.84**	-	-	1		
Non-experts (2)	-	0.77**	-	0.50**	1	
Non-experts (3)	-	-	0.66**	0.46**	0.92**	1

**p<0.01

In the non-expert group, the strongest relationship occurs between Conditions 2 and 3, where the subjects were aware of the type of information that should be taken into account in rating but no actual information was given. Condition 1, where the most information was provided, is only slightly more strongly related to Condition 2 than to Condition 3. This suggests that the amount of information subjects were given made the most difference in Condition 1, rather than the matrix approach itself. In the expert group, Condition 1 has a stronger relationship with Condition 2 than Condition 3, while Condition 3 has a stronger relationship with Condition 2 than it does with Condition 1. This reflects the amount of information that subjects were given in each condition.

The strongest relationship between expert and non-expert ratings is in Condition 1 where most information was given to both groups. Condition 2 shows a weaker relationship and Condition 3 shows the weakest of all.

The relative rankings were compared with the rankings derived in Chapter 4 using the risk assessment procedure (see Table 9). All rankings derived from the expert subject group in Condition 1 are within 5 places of the original ranking, with the exception of "sulphur dioxide and smoke (domestic)" (11 places), "house dust mites (non-domestic)" (9 places), "slips, trips and falls from buildings (non-domestic)" (8 places) and "noise (domestic)" (7 places).

All the rankings derived from the expert subject group under the 3 conditions correlate at a significant level with the original risk assessment procedure ranking, with a coefficient increasing with the amount of information given. The same is true for non-expert rankings, although these

are of a lesser magnitude and those derived from Condition 3 do not give a significant correlation.

5.2.4 Conclusion

Once they are in possession of risk information, in the form of risk matrices, non-experts are as similar to experts in possession of the same information as two groups of experts without that information, i.e. experts in Conditions 2 and 3. This implies that the reasons experts and non-experts assess risks differently stem from differences in knowledge, rather than differences in combining different pieces of information (at least when no personal interest is involved).

While the judgements made in this study cannot be used, strictly speaking, to validate the risk assessment methodology, they do increase confidence in the procedure itself and in the degree of acceptance that its results are likely to achieve.

5.3 Numerical weighting of severity of harm

5.3.1 Introduction

The BRE risk assessment methodology classifies harms in Classes I to IV, with weightings assigned to each Class, for the purpose of calculating a risk index. It was found, in a sensitivity analysis of these weightings, that hazards causing large numbers of less serious harms tended to move in relation to hazards causing a small number of more serious harms, resulting in uncertainty in the use of ranking. Additionally, no attempt had been made to assign different weights to different lives according to factors such as the age of the person affected. This study therefore aimed to increase confidence in the weightings given to Classes I to IV and to obtain relative weightings for different age groups. The work is described in more detail in Appendix G.

5.3.2 Method

Subjects were shown a table which illustrated the four different classes of harm and were asked to assign values to these classes of harm, given that one million pounds would avoid ten cases of the most severe harms (Class I harms). Subjects were asked to decide how many harms in the other classes should be avoided by spending the same amount of money as was spent to avoid the 10 Class I harms. Once this part of the task had been completed, the subjects were asked to repeat the same procedure but for different age groups (under 18, 18-40, 40-60 and over 60).

There was concern that subjects asked to weight classes of harm did not base their assessments on a detailed evaluation of the harms, but rather adopted a nominal multiplication factor between classes. Therefore, in a supplementary study, expert subjects gave weightings for six classes of harm, rather than four. The six classes were devised by splitting each of the original Classes II and III into two classes, so that Classes I and IV retained the same meaning but had more Classes intervening between them.

5.3.3 Results

Medians for each age group and each class of harm are shown in Table 11. Experts' weightings increased with the age of the persons affected while non-experts' weightings decreased with age. These opposing approaches to the treatment of age makes it difficult to make any recommendations for the inclusion of age as a factor in the risk assessment methodology, especially given the relatively small magnitude of the age effects.

The weightings differed between the three measures of central tendency used (arithmetic mean, geometric mean and median). For example the ratio between Class I and Class IV could be 50 or

10000, depending on which results are regarded as most valid and reliable. The median is the preferred measure because the distributions are highly skewed and sometimes bimodal.

Table 11. Medians of severity weightings, normalised to Class I

AGE OF PERSONS AFFECTED	SUBJECT GROUP	HARM CLASS			
		IV	III	II	I
Not specified	Experts	1000	28	10	1
	Non-Experts	50	11	5	1
Under 18	Experts	1000	50	10	1
	Non-Experts	60	30	7.5	1
18 - 40	Experts	1000	25	10	1
	Non-Experts	25	12.5	5	1
40 - 60	Experts	1000	32	10	1
	Non-Experts	30	10	5	1
Over 60	Experts	1000	20	10	1
	Non-Experts	20	12	4	1

When the number of classes was increased to six, the relationship between Classes I and II was relatively stable, with a little more variation in Classes III and IV.

5.3.4 Conclusion

Comparison of expert and non-expert severity weightings showed that there were differences between the two groups. None of the figures comes close to the 100,000 used in the original risk assessment methodology.

The expert medians in the main study do not vary with age group, except in the case of Class III, for which the weighting varies from 20 to 50 (median of medians 35, mean of medians 31). It is also for Class III that there is least consensus between studies, although the variation is not of such a magnitude that it is likely to have a major impact on risk assessment outcomes. Giving greater weight to the subjects in the main study (who carried out a range of tasks and therefore had more chance to become familiar with the concepts), on balance a figure of 30 is suggested for Class III.

Putting this together with the other classes indicates values of 1, 10, 30 and 1000 for Classes IV (least severe) to I (most severe) respectively.

Subjects on the whole did not greatly alter weightings when applying them to different age groups, suggesting that they did not value lives differently for different age groups.

5.4 Numerical weighting of strength of evidence

5.4.1 Introduction

The risk assessment methodology assigns numerical weightings to the strength of the evidence for a link between a hazard and the number of people affected by a given class of harm (see Chapter 3). The strength of evidence was classified as follows:

High (***) a link between hazard and outcome is well established by firm evidence;

Medium (**) probably an effect but less complete evidence is available;

Low (*) possibly a link between hazard and outcome.

These strengths of evidence were assigned values of 5, 3 and 1 respectively in the calculation of a risk index. The aim of the study was to increase confidence in these values by using costs of avoidance as an index, using judgements made by experts and non-experts. The work is described in more detail in Appendix H.

5.4.2 Method

The method was essentially the same procedure as the severity weighting study except that the three levels of strength of evidence were compared and age groups were not separated. The subjects were then asked to assign a value to each level of evidence for each class of harm, given that one million pounds (£1,000,000) would certainly avoid a very severe harm (Class I). Thus, the larger the number given by the subjects, the more severe the combination of evidence and class of harm was considered to be.

5.4.3 Results and discussion

Table 12 shows the median weightings for each level of strength of evidence and each class of harm for both expert and non-expert subjects. The scores are normalised to a value of 1 for low strength of evidence in each case; therefore comparisons cannot be made between classes.

5.4.4 Conclusion

Weightings given to strength of evidence were more consistent than those for class of harm. The medians from the expert subject group may be the most appropriate to adopt for the high, medium and low values of strength of evidence since the medians are very similar across the four classes. These medians would suggest that values of 1, 2 and 4 be adopted for low, medium and high strengths of evidence compared to 1, 3 and 5 used in the original risk index methodology.

Table 12. Median weightings given by experts and non-experts

	EXPERTS			NON-EXPERTS		
STRENGTH OF EVIDENCE →	HIGH	MEDIUM	LOW	HIGH	MEDIUM	LOW
Harm Class IV	4.5	2.3	1	3.2	1.8	1
Harm Class III	4.5	2.0	1	3.3	1.9	1
Harm Class II	4.0	2.0	1	2.6	1.5	1
Harm Class I	4.0	2.2	1	2.0	1.6	1

5.5 Assessment of the most severe harms

5.5.1 Introduction

In Chapter 2, subjects rated the severity of a list of harms directly against a seven point scale, from 1 (extremely slight) to 7 (extremely severe). The study resulted in 28 harms being placed within the 6-7 band. The aim of the present study was to rate the harms in this most severe band in order to give a more precise comparison of harms within this band and thus provide greater confidence in stating the degrees of harm associated with specific health outcomes. In particular, the high degree of agreement found between experts and non-experts might be attributed to the major differences among the harms rated. If this is the case, then a lesser degree of agreement should be found when a narrower range of harms is assessed. The work is described in more detail in Appendix I.

5.5.2 Method

The subjects were shown a list of 38 harms in the 6-7 rating band (i.e. the most serious), and

asked to rate each of the harms against a seven point scale in which 1 equalled “least severe” and 7 equalled “most severe”. The method was otherwise as in Chapter 2, that is, to sort through the harms and place them in at the appropriate point on the scale.

5.5.3 Results

There was a strong linear relationship between expert and non-expert ratings of severity of harm (r=0.94) and, of the 38 harms, only four showed significant differences between expert and non-expert ratings (see Appendix I for details). Fatal harms were rated as the most severe and the lower end of the severity scale tended to be avoided

5.5.4 Conclusion

The high degree of agreement between subject groups supports the use of the severity scale as derived in Chapter 2.

5.6 Impact on risk methodology

The values for weightings of severity of harm and strength of evidence derived above can be substituted into the risk index methodology to produce new risk indices for all hazards. The strength of evidence weightings were not greatly affected by the current study, and the change was within that shown by sensitivity analysis (Chapter 4) to have little effect on the risk index. Any effects must therefore be due largely to the change in harm class weightings for class of harm. The rank order produced by these new indices has been correlated with the original index and with the hazard ratings given by experts and non-experts in Appendix F (see Table 13). The table also shows correlations between the hazard ratings and the original risk index.

Table 13. Rank correlation coefficients between hazard ratings and risk indices

	ORIGINAL RISK INDEX	EXPERT RATINGS			NON-EXPERT RATINGS		
		1	2	3	1	2	3
New risk index	0.860**	0.723**	0.485*	0.348	0.672**	0.070	0.041
Original risk index	1	0.917**	0.732**	0.662**	0.690**	0.366*	0.304

* p< 0.05 ** p< 0.01

It is interesting to note that the correlations produced with the ratings with the new risk index are lower than those with the original risk index. This suggests that subjects may be placing more emphasis on deaths and the most severe harms when considering the hazards than when considering weighting severity on its own.

The new risk indices do change the overall rankings of all hazards and would give a somewhat different complexion to the rank order groupings shown Chapter 4, based on the original risk index methodology. Table 14 shows health and safety hazards grouped in rank order of risk, based on the new values for severity weighting and strength of evidence. The hazards that have moved levels in the list are indicated.

Around half of the domestic hazards (14 from 29) change levels, although only one (slips trips and falls involving baths) moves more than one level. Those in the top level of risk are quite similar to the original ranking, which confirms that these require the most attention. Over half of the non-domestic hazards change levels (14 from 26) but none by more than one.

Table 14. Health and safety hazards, grouped in rank order of risk

HOMES	NON-DOMESTIC BUILDINGS
Highest risk	
Hygrothermal conditions Radon Slips, trips and falls on stairs and ramps Burns/scald hazards Fire hazards House dust mites Environmental tobacco smoke Noise[+]	
Second level of risk	
Fungal growth Slips, trips and falls on the level[-] Carbon monoxide[-] Lead Slips, trips and falls from buildings[-] Drowning hazards[-] Oxides of nitrogen[+] Sanitary accommodation[+] Electrical hazards	Slips, trips and falls on the level[-] Environmental tobacco smoke Hygrothermal conditions[+] Lighting[+] Space[+] House dust mites[+] Slips, trips and falls on stairs and ramps Sources of infection other than sanitary accommodation Radon[-]
Third level of risk	
Sources of infection other than sanitary accommodation Volatile organic compounds Slips, trips and falls involving baths[- -] Architectural glass[-] Explosions[-] Collision/entrapment involving doors Collision/entrapment involving windows[-] Sulphur dioxide and smoke[+]	Volatile organic compounds Slips, trips and falls from buildings[-] Electrical hazards[-] Fungal growth Noise[+] Explosions[-] Carbon monoxide[-] Sanitary accommodation[+]
Fourth level of risk	
Particulates[-] Falling objects[-] Landfill gas Space[-] Pesticides	Oxides of nitrogen Falling objects[-] Collision/entrapment involving doors[-] Collision/entrapment involving windows Collision/entrapment involving lifts

[-] indicates the hazard has moved down a level from the original rank
[+] indicates the hazard has moved up a level from the original rank

6. CONCLUSION

The work reported here has produced and applied an initial procedure for comparing the risks associated with specified hazards in buildings. The procedure is applicable to a wide range of hazards. The risk index has proved very useful in generating a rank order of hazards, but its limitations have to be recognised, as discussed above; in summary:

- the procedure does not assign different weights to different lives (e.g. according to age);
- the rank ordering does not take into account possible justifications for action other than health;
- the procedure does not indicate where current legislation and practice are containing what would otherwise be a greater risk;
- the ranks given will depend on the grouping of hazards adopted in the reports;
- a low rank may mean that only poor quality data or no data at all are available.

The main area of uncertainty is the weighting of severity of harm. For some purposes it may, therefore, be appropriate to calculate indices using both the original and the new weightings. This would indicate the range of ranks that might be given to each hazard.

REFERENCES

Cox, T & Cox, S J (1992). *Occupational Health: The Control and Monitoring of and Organisational Hazards in the Workplace*. Report to the World Health Organisation, European Office, Copenhagen, Denmark.

Cox, S J & O'Sullivan, E (1995) *Building regulation and safety.* BRE Report. Garston: Construction Research Communications.

Cox, S J & Smithies, N (1998) *Building regulation and fire safety.* Draft BRE Report.

Cox, S J & Tait, N R S (1991) *Reliability, Safety and Risk Management.* Oxford: Butterworth-Heinemann.

Keppel, G, Saufley, W H & Tokunga, H (1992) Introduction to design and analysis (2nd ed.). WH Freeman & Co, New York.

Mant, D & Gray, J M (1986) *Building regulation and health*. BRE Report. Garston: Construction Research Communications.

Raw, G J & Hamilton, R M (1995) *Building regulation and health.* BRE Report. Garston: Construction Research Communications.

APPENDIX A.
GENERATION OF A LIST OF REPRESENTATIVE HARMS

A1. INTRODUCTION

This first stage of the work was intended to identify the most probable types of harm associated with building hazards, and to classify those harms, with the purpose of producing a representative list of harms.

A2. DATA COLLECTION

It was assumed that the ongoing BRE projects on health, safety and fire safety in buildings involved a representative sample of expert opinion on building hazards and associated types of harm. A group of 12 experts was drawn from those involved in these projects and was challenged to identify the most probable types of harm associated with building hazards.

Brainstorming techniques were used with this group of experts to elicit a list of harms. This technique was developed by Osborn in the 1930s for eliciting a large number of ideas from a small group of people in a short period of time; it is commonly governed by the following rules:

- the group facilitator has to provide an adequate description of the task and its context;
- participants should not be afraid to contribute and everyone should contribute, without criticism between members of the group;
- the free flow of ideas should be encouraged and as many ideas should be recorded as possible;
- participants should attend to developing a written record of ideas, which must be clearly visible;
- participants should be creative and far fetched ideas should not be discouraged.

The group facilitator started by describing the task in hand. Participants then recorded their ideas concisely on cards, which were periodically collected and pinned to a board. Each idea was read out as it was pinned up, to provide a constant background of ideas for the continuing generation of new ideas. This process was repeated until it was agreed that all the most important probable harms had been identified.

In order to create a degree of consistency in the format of the ideas, participants were issued with simple rules on levels of description. These rules stated that each statement should contain a general anatomical location (e.g. arm, torso, face) and the general nature of the harm (e.g. broken bone, bruise). It was recognised that the harm sometimes implies the location (e.g. pneumonia implies lungs).

The group was then split into four subgroups, each containing one member from each field (health, safety, fire safety). Each subgroup was challenged to identify the minimum number of exclusive (and independent) categories which could then be used to describe the standardised data. Each group was interviewed after this exercise about the logic underpinning its classification.

A3. CLASSIFICATION OF HARMS

The concepts used by the group in classifying the harms can be reduced to (i) anatomical location, (ii) severity, (iii) the cause of the harm and (iv) the duration or time-course of the harm. Of these, the cause of harm is not useful for our purposes since a cause-related classification cannot be applied equally over a range of hazards.

Building hazards appeared to fulfil one of two possible roles in relation to harms: first, it might cause or directly contribute to the cause of those harms (aetiological role) or second, it might exacerbate the harm (moderator role).

A4. TREATMENT OF THE LIST OF HARMS

The list of harms and the structures generated by the subgroups were reviewed several times and a working list of harms produced.

First the number of harms was reduced, giving careful consideration to:

- redundancy (repetition) in the list:
- variation in level of specificity in the descriptions:
- inclusion of items that were not harms (e.g. they were hazards, or groups of harms).

The original list of 308 items was thus reduced by 47% to 164.

The next stage involved expanding the list of 164 harms to cover a wider range of harms and qualifying those harms in a consistent manner.

Most of the harms were found to be focused on safety and fire safety. A draft of the BRE report,

Building regulation and health, was therefore used as a basis for the inclusion of additional health-related harms to balance the list.

Two types of qualifier were then chosen for inclusion in the listing of harms, timing of occurrence and severity; these qualifiers reflected the thinking of the expert group. The qualifiers were applied to every harm to which they logically could be applied, not just those to which the experts happened to have applied them in the brainstorming session. In many cases the immediacy of the harm is implied by the nature of the harm (e.g. broken arm) and is therefore not specified. Similarly not every level of severity is applicable to every harm, and the same qualifier can have a different meaning for different harms (e.g. 'severe' does not necessarily carry the same meaning whether applied to eyestrain or to severe multiple bruising).

In addition, each harm was put in the context of the full range of plausible anatomical locations, the other main classification variable used by the expert group.

The result of this expansion was a list of 748 harms (Appendix B).

Finally the list was reduced by 60% by allowing only a representative selection of qualifiers for each harm, representative that is of the full range of severity which each harm might represent. The final list of 297 harms is shown in bold type in Appendix B.

A5. CONCLUSIONS

The first phase worked reasonably well although the expert group did not operate exactly as expected. This appeared to be partly a reflection of the strength of members' existing concepts concerning hazards and harms, and the frameworks within which they normally operated.

Four difficulties arose with the data they generated: (a) there was some confusion in the data over what was meant by harm, (b) the group did not focus on the most probable harms, (c) the group focused mainly on harms relating to safety and fire safety, and (d) the group could not agree on a common level of description for the harms generated.

The subsequent treatment of the list of harms made it possible to plan the second stage, an analysis of the conceptualisation of harm and overall rating of degree of harm. This is reported in Appendix C.

ACKNOWLEDGEMENTS

The authors are grateful for the assistance of Dr Ruth Hamilton for medical advice on changes to the list of harms, and Catherine Horne for producing the sorting cards and the first list of harms.

APPENDIX B.
FINAL LIST OF HARMS FOR SEVERITY RATING

This is the list of harms generated by the procedure described in Appendix A. The shorter list of representative harms is in bold text. The descriptions of the harms do not always correspond to exact medical terminology (e.g. 'cancer' implies malignancy but the word 'malignant' is included in the description of cancers to emphasise that benign tumours are excluded).

Addiction to legal drugs (e.g. alcohol, nicotine, prescribed drugs)
Use of class 1 illegal drugs
Use of illegal drugs (other than class 1)

Occasional irritability/intolerance towards other occupants
Occasional aggressive behaviour when in building
Occupant occasionally violent when in building
Occupant becomes occasionally violent over time
Regular irritability/intolerance towards other occupants
Regular aggressive behaviour when in building
Occupant regularly violent when in building
Occupant becomes regularly violent over time

Alienation from society (as an immediate result of occupying a certain home)
Alienation from society (over time because of changes to self)

Short term amnesia
Long term amnesia

Occasional slight annoyance
Occasional moderate annoyance
Occasional serious annoyance
Regular slight annoyance
Regular moderate annoyance
Regular serious annoyance
Constant slight annoyance
Constant moderate annoyance
Constant serious annoyance

Occasional slight anxiety
Occasional moderate anxiety
Occasional serious anxiety
Regular slight anxiety
Regular moderate anxiety
Regular serious anxiety
Constant slight anxiety
Constant moderate anxiety
Constant serious anxiety

Onset of mild asthmatic condition
Onset of serious asthmatic condition

Asphyxiation with full recovery
Asphyxiation with minor brain damage
Asphyxiation with major brain damage
Asphyxiation resulting in death

Temporary slight backache
Temporary moderate backache
Temporary severe backache
Chronic slight backache
Chronic moderate backache
Chronic severe backache

Temporary loss of hair in accident
Temporary loss of hair because of disease
Temporary loss of hair because of burns
Permanent loss of hair in accident
Permanent loss of hair because of disease
Permanent loss of hair because of burns

Behaviour disorder:
mild temporary (immediately following hazardous event)
mild chronic (immediately following hazardous event)
mild temporary (developing over time)
mild chronic (developing over time)
moderate temporary (immediately following hazardous event)
moderate chronic (immediately following hazardous event)
moderate temporary (developing over time)
moderate chronic (developing over time)
serious temporary (immediately following hazardous event)
serious chronic (immediately following hazardous event)
serious temporary (developing over time)
serious chronic (developing over time)

Benign tumour(s)
Malignant skin cancer - treatable with complete cure
Malignant lung cancer - treatable with complete cure
Other malignant cancer - treatable with complete cure
Leukaemia - treatable with complete cure
Malignant skin cancer - treatable with risk of recurrence or development of secondaries
Malignant lung cancer - treatable with risk of recurrence or development of secondaries
Other malignant cancer - treatable with risk of recurrence or development of secondaries
Leukaemia - treatable with risk of recurrence or development of secondaries
Malignant skin cancer - non-treatable
Malignant lung cancer - non-treatable
Other malignant cancer - non-treatable
Leukaemia - non-treatable

Loss of:
arm in accident
arm by amputation because of injury
arm by amputation because of disease
arm because of burns
arms in accident
arms by amputation because of injury
arms by amputation because of disease
arms because of burns
arm because of burns
arms in accident
arms by amputation because of injury
arms by amputation because of disease
arms because of burns
arms & legs in accident
arms & legs by amputation because of injury
arms & legs by amputation because of disease
arms & legs because of burns
dominant hand in accident
dominant hand by amputation because of injury
dominant hand by amputation because of disease
dominant hand because of burns
non-dominant hand in accident
non-dominant hand by amputation because of injury
non-dominant hand by amputation because of disease
non-dominant hand because of burns
finger in accident
finger by amputation because of injury
finger by amputation because of disease
finger because of burns
thumb in accident
thumb in accident
thumb by amputation because of injury
thumb by amputation because of disease
thumb because of burns
finger/thumb/toe nail in accident
finger/thumb/toe nail by surgical removal
leg in accident
leg by amputation because of injury
leg by amputation because of disease
leg because of burns
legs in accident
legs by amputation because of injury
legs by amputation because of disease
legs because of burns
foot in accident
foot by amputation because of injury
foot by amputation because of disease
foot because of burns
feet in accident
feet by amputation because of injury
feet by amputation because of disease
feet because of burns
toes in accident
toes by amputation because of injury
toes by amputation because of disease
toes because of burns

Occasional insect bites
Frequent insect bites

Immediate temporary impaired vision in one eye
Immediate temporary impaired vision in both eyes
Immediate permanent impaired vision in one eye
Immediate permanent impaired vision in both eyes
Gradual temporary impaired vision in one eye
Gradual temporary impaired vision in both eyes
Gradual permanent impaired vision in one eye
Gradual permanent impaired vision in both eyes
Immediate temporary blindness in one eye
Immediate temporary blindness in both eyes
Immediate permanent blindness in one eye
Immediate permanent blindness in both eyes
Gradual temporary blindness in one eye
Gradual temporary blindness in both eyes
Gradual permanent blindness in one eye
Gradual permanent blindness in both eyes

Mild claustrophobia
Moderate claustrophobia
Severe claustrophobia

Minor immediate brain damage
Moderate immediate brain damage
Serious immediate brain damage
Fatal immediate brain damage
Minor gradual brain damage
Moderate gradual brain damage
Serious gradual brain damage
Fatal gradual brain damage

Moderate bruising to feet/ankles
Severe bruising to feet/ankles
Moderate bruising to arms/wrists
Severe bruising to arms/wrists
Moderate bruising to face
Severe bruising to face
Moderate bruising to hands
Severe bruising to hands
Moderate bruising to legs
Severe bruising to legs
Moderate bruising to head/neck
Severe bruising to head/neck
Moderate bruising to torso
Severe bruising to torso
Moderate multiple bruising
Severe multiple bruising

Superficial blisters on face
Superficial blisters on hands
Superficial blisters other than on face or hands
Serious blisters on face
Serious blisters on hands
Serious blisters other than on face or hands

Minor burns/scalds to legs/ankle
Moderate burns/scalds to legs/ankle

Severe burns/scalds to legs/ankle
Minor burns/scalds to arms/wrists
Moderate burns/scalds to arms/wrists
Severe burns/scalds to arms/wrists
Minor burns/scalds to face
Moderate burns/scalds to face
Severe burns/scalds to face
Minor burns/scalds to feet
Moderate burns/scalds to feet
Severe burns/scalds to feet
Minor burns/scalds to hands
Moderate burns/scalds to hands
Severe burns/scalds to hands
Minor burns/scalds to head/neck
Moderate burns/scalds to head/neck
Severe burns/scalds to head/neck
Minor burns/scalds to torso
Moderate burns/scalds to torso
Severe burns/scalds to torso
Minor burns/scalds to wrists
10% burns/scalds
20% burns/scalds
40% burns/scalds
60% burns/scalds
80% burns/scalds

Slight concussion
Moderate concussion
Severe concussion

Acute moderate bronchitis
Acute severe bronchitis
Acute fatal bronchitis
Chronic moderate bronchitis
Chronic severe bronchitis
Chronic fatal bronchitis

Occasional moderate diarrhoea/vomiting
Occasional severe diarrhoea/vomiting
Regular moderate diarrhoea/vomiting
Regular severe diarrhoea/vomiting

Wheezing/breathlessness:
temporary slight
temporary moderate acute
chronic slight (immediately induced)
chronic moderate (immediately induced)
chronic severe (immediately induced)
chronic slight (gradually induced)
chronic moderate (gradually induced)
chronic severe (gradually induced)

Mild temporary confusion/disorientation
Moderate temporary confusion/disorientation
Severe temporary confusion/delerium
Mild chronic confusion/disorientation
Moderate chronic confusion/disorientation
Severe chronic confusion/dementia

Minor dislocated ankle
Minor dislocated toe
Minor dislocated spine
Minor dislocated elbow
Minor dislocated finger
Minor dislocated hip
Minor dislocated jaw
Minor dislocated knee
Minor dislocated shoulder
Minor dislocated wrist
Serious dislocated ankle
Serious dislocated toe
Serious dislocated spine
Serious dislocated elbow
Serious dislocated finger
Serious dislocated hip
Serious dislocated jaw
Serious dislocated knee
Serious dislocated shoulder
Serious dislocated wrist

Subclinical short-term depression
Mild short-term depression
Severe short-term depression
Subclinical long-term depression
Mild long-term depression
Severe long-term depression

Occasional slight dermatitis
Occasional moderate dermatitis
Occasional severe dermatitis
Regular slight dermatitis
Regular moderate dermatitis
Regular severe dermatitis

Occasional slight discomfort
Occasional moderate discomfort
Occasional severe discomfort
Regular slight discomfort
Regular moderate discomfort
Regular severe discomfort

Occasional slight dissatisfaction
Occasional moderate dissatisfaction
Occasional serious dissatisfaction
Regular slight dissatisfaction
Regular moderate dissatisfaction
Regular serious dissatisfaction

Occasional immediate slight drowsiness
Occasional immediate moderate drowsiness
Occasional immediate severe drowsiness
Regular immediate slight drowsiness
Regular immediate moderate drowsiness
Regular immediate severe drowsiness
Occasional delayed slight drowsiness
Occasional delayed moderate drowsiness
Occasional delayed severe drowsiness
Regular delayed slight drowsiness
Regular delayed moderate drowsiness
Regular delayed severe drowsiness

Slight electric shock
Moderate electric shock
Severe electric shock
Fatal electric shock

Occasional slight eye strain
Occasional moderate eye strain
Occasional severe eye strain

Regular slight eye strain
Regular moderate eye strain
Regular severe eye strain

Brief loss of consciousness
Frequent brief loss of consciousness
Loss of consciousness for hours/days
Frequent loss of consciousness for hours/days
Long term loss of consciousness
Permanent loss of consciousness

Occasional slight fatigue
Occasional moderate fatigue
Occasional severe fatigue/exhaustion
Regular slight fatigue
Regular moderate fatigue
Regular severe fatigue/exhaustion

Fear of real threat (e.g. crime, children's safety):
occasional slight
occasional moderate
occasional severe
regular slight
regular moderate
regular severe

Irrational fear (e.g. of spiders, enclosed spaces):
occasional slight
occasional moderate
occasional severe
regular slight
regular moderate
regular severe

Occasional slight fever
Occasional moderate fever
Occasional severe fever
Regular slight fever
Regular moderate fever
Regular severe fever

Occasional petit mal seizure
Regular petit mal seizure
Occasional grand mal seizure
Regular grand mal seizure

Gangrene in finger/thumb
Gangrene in hand
Gangrene in wrist/lower arm
Gangrene in toe
Gangrene in foot
Gangrene in ankle/lower leg
Gangrene in upper limb

Stomach/duodenal ulcer

Slight graze to arm/leg
Slight graze to hand
Slight graze to face
Slight graze to body
Bad graze to arm/leg
Bad graze to hand
Bad graze to face
Bad graze to body

Occasional slight headache
Occasional moderate headache
Occasional severe headache
Regular slight headache
Regular moderate headache
Regular severe headache

Moderate internal abdominal damage
Serious internal abdominal damage
Fatal internal abdominal damage

Immediate mild heart attack
Immediate moderate heart attack
Immediate severe heart attack
Immediate fatal heart attack
Delayed mild heart attack
Delayed moderate heart attack
Delayed severe heart attack
Delayed fatal heart attack

Short-term slight high blood pressure
Short-term moderate high blood pressure
Short-term severe high blood pressure
Chronic slight high blood pressure
Chronic moderate high blood pressure
Chronic severe high blood pressure

Occasional slight impairment of children's play
Occasional moderate impairment of children's play
Occasional severe impairment of children's play
Regular slight impairment of children's play
Regular moderate impairment of children's play
Regular severe impairment of children's play

Occasional slight impairment of concentration
Occasional moderate impairment of concentration
Occasional severe impairment of concentration
Regular slight impairment of concentration
Regular moderate impairment of concentration
Regular severe impairment of concentration

***Fracture**:*
hairline, of ankles/feet
simple, of ankles/feet
compound, complicated or comminuted, of ankles/feet
hairline, of arms/shoulder
simple, of arms/shoulder
compound, complicated or comminuted, of arms/shoulder
hairline, of back
simple, of back without damage to spinal cord
fracture of back with damage to spinal cord
hairline, of wrist/hands
simple, of wrist/hands
compound, complicated or comminuted, of wrist/hands

hairline, of legs
simple, of legs
compound, complicated or comminuted, of legs
hairline, of neck
of neck without damage to spinal cord
fracture of neck with damage to spinal cord
hairline, of pelvis
simple, of pelvis
compound, complicated or comminuted, of pelvis
hairline, of ribs
simple, of ribs
compound, complicated or comminuted, of ribs
hairline, of skull
simple, of skull
compound, complicated or comminuted, of skull
hairline, of fingers
simple, of fingers
compound, complicated or comminuted, of fingers
hairline, of toes
simple, of toes
compound, complicated or comminuted, of toes
hairline, of nose
simple, of nose
compound, complicated or comminuted, of nose
hairline, of jaw
simple, of jaw
compound, complicated or comminuted, of jaw
multiple simple fractures
multiple compound, complicated or comminuted fractures

Immediate temporary impaired hearing in one ear
Immediate temporary impaired hearing in both ears
Immediate permanent impaired hearing in one ear
Immediate permanent impaired hearing in both ears
Gradual temporary impaired hearing in one ear
Gradual temporary impaired hearing in both ears
Gradual permanent impaired hearing in one ear
Gradual permanent impaired hearing in both ears
Immediate temporary deafness in one ear
Immediate temporary deafness in both ears
Immediate permanent deafness in one ear
Immediate permanent deafness in both ears
Gradual temporary deafness in one ear
Gradual temporary deafness in both ears
Gradual permanent deafness in one ear
Gradual permanent deafness in both ears

Impairment of judgement:
occasional/temporary slight
occasional/temporary moderate
occasional/temporary severe
regular/chronic slight
regular/chronic moderate
regular/chronic severe

Mild hyperthermia
Moderate hyperthermia
Severe hyperthermia
Fatal hyperthermia
Mild hypothermia
Moderate hypothermia
Severe hypothermia
Fatal hypothermia

Impairment of sexual behaviour:
occasional/temporary slight
occasional/temporary moderate
occasional/temporary severe
regular/chronic slight
regular/chronic moderate
regular/chronic severe

Minor hernia
Major hernia

Occasional slight restriction of social behaviour
Occasional moderate restriction of social behaviour
Occasional severe restriction of social behaviour
Regular slight restriction of social behaviour
Regular moderate restriction of social behaviour
Regular severe restriction of social behaviour

Mild angina
Sever angina

Temporary/occasional slight irritation of eyes
Temporary/occasional moderate irritation of eyes
Temporary/occasional severe irritation of eyes
Chronic/regular slight irritation of eyes
Chronic/regular moderate irritation of eyes
Chronic/regular severe irritation of eyes

Temporary/occasional slight irritation of nose
Temporary/occasional moderate irritation of nose
Temporary/occasional severe irritation of nose
Chronic/regular slight irritation of nose
Chronic/regular moderate irritation of nose
Chronic/regular severe irritation of nose

Temporary/occasional slight irritation of throat
Temporary/occasional moderate irritation of throat
Temporary/occasional severe irritation of throat
Chronic/regular slight irritation of throat
Chronic/regular moderate irritation of throat
Chronic/regular severe irritation of throat

Temporary/occasional slight irritation of skin
Temporary/occasional moderate irritation of skin
Temporary/occasional severe irritation of skin
Chronic/regular slight irritation of skin
Chronic/regular moderate irritation of skin
Chronic/regular severe irritation of skin
Temporary/occasional redness/rash on face
Chronic/regular redness/rash on face

Temporary/occasional redness/rash on hands
Chronic/regular redness/rash on hands
Temporary/occasional redness/rash other than on face or hands
Chronic/regular redness/rash other than on face or hands

Occasional mild cough/cold
Occasional moderate cough/cold
Occasional serious cough/cold
Regular mild cough/cold
Regular moderate cough/cold
Regular serious cough/cold

Occasional influenza
Regular influenza

Occasional slight lack of privacy
Occasional moderate lack of privacy
Occasional severe lack of privacy
Regular slight lack of privacy
Regular moderate lack of privacy
Regular severe lack of privacy

Occasional slight loss of appetite
Occasional moderate loss of appetite
Occasional severe loss of appetite
Regular slight loss of appetite
Regular moderate loss of appetite
Regular severe loss of appetite

Slightly low birth weight
Very low birth weight

Occasional slight menstrual disorder
Occasional moderate menstrual disorder
Occasional severe menstrual disorder
Regular slight menstrual disorder
Regular moderate menstrual disorder
Regular severe menstrual disorder

Temporary/occasional slight muscular weakness
Temporary/occasional moderate muscular weakness
Temporary/occasional severe muscular weakness
Chronic/regular slight muscular weakness
Chronic/regular moderate muscular weakness
Chronic/regular severe muscular weakness

Restricted child development (social/emotional/cognitive):
slight temporarily
moderate temporarily
severe temporarily
slight permanently
moderate permanently
severe permanently

Occasional slight eye strain
Occasional moderate eye strain
Occasional severe eye strain
Regular slight eye strain
Regular moderate eye strain
Regular severe eye strain

Temporary paralysis of foot/feet
Temporary paralysis of leg
Temporary paralysis of legs
Temporary paralysis of arm
Temporary paralysis of arms
Temporary paralysis below waist
Temporary paralysis below neck
Permanent paralysis of foot/feet
Permanent paralysis of leg
Permanent paralysis of legs
Permanent paralysis of arm
Permanent paralysis of arms
Permanent paralysis below waist
Permanent paralysis below neck

Occasional slight rhinitis
Occasional moderate rhinitis
Occasional severe rhinitis
Regular slight rhinitis
Regular moderate rhinitis
Regular severe rhinitis

Temporary immediate slight personality change
Temporary immediate moderate personality change
Temporary immediate severe personality change
Chronic immediate slight personality change
Chronic immediate moderate personality change
Chronic immediate severe personality change
Temporary delayed slight personality change
Temporary delayed moderate personality change
Temporary delayed severe personality change
Chronic delayed slight personality change
Chronic delayed moderate personality change
Chronic delayed severe personality change

Occasional moderate migraine
Occasional severe migraine
Regular moderate migraine
Regular severe migraine

Occasional mild pneumonia
Occasional moderate pneumonia
Occasional severe pneumonia
Regular mild pneumonia
Regular moderate pneumonia
Regular severe pneumonia

Temporary slight psychosis
Temporary moderate psychosis
Temporary severe psychosis
Chronic slight psychosis
Chronic moderate psychosis
Chronic severe psychosis

Minor puncture/cuts to ankles/feet
Moderate puncture/cuts to ankles/feet
Serious puncture/cuts to ankles/feet
Minor puncture/cuts to arm/wrist
Moderate puncture/cuts to arm/wrist
Serious puncture/cuts to arm/wrist

Minor puncture/cuts hands
Moderate puncture/cuts hands
Serious puncture/cuts hands
Minor puncture/cuts to head/face
Moderate puncture/cuts to head/face
Serious puncture/cuts to head/face
Minor puncture/cuts to leg
Moderate puncture/cuts to leg
Serious puncture/cuts to leg
Minor puncture/cuts to neck
Moderate puncture/cuts to neck
Serious puncture/cuts to neck
Minor puncture/cuts to torso
Moderate puncture/cuts to torso
Serious puncture/cuts to torso
Multiple minor punctures/cuts
Multiple moderate punctures/cuts
Multiple serious punctures/cuts

Slight birth defect
Moderate birth defect
Severe birth defect
Fatal birth defect

Slight retinal haemorrhage
Moderate retinal haemorrhage
Severe retinal haemorrhage

Slight immunological sensitisation to chemical/biological agents
Moderate immunological sensitisation to chemical/biological agents
Severe immunological sensitisation to chemical/biological agents

Temporary mild sleep disorder
Temporary moderate sleep disorder
Temporary serious sleep disorder
Chronic mild sleep disorder
Chronic moderate sleep disorder
Chronic serious sleep disorder

Occasional mild sleep deprivation
Occasional moderate sleep deprivation
Occasional serious sleep deprivation
Regular mild sleep deprivation
Regular moderate sleep deprivation
Regular serious sleep deprivation

Immediate temporary partial dependency
Immediate temporary complete dependency
Immediate permanent partial dependency
Immediate permanent complete dependency
Gradual temporary partial dependency
Gradual temporary complete dependency
Gradual permanent partial dependency
Gradual permanent complete dependency

Temporary slight speech impairment
Temporary moderate speech impairment
Temporary severe speech impairment
Chronic slight speech impairment
Chronic moderate speech impairment
Chronic severe speech impairment

Slight slipped disc
Moderate slipped disc
Serious slipped disc

Slight ligament injury in ankle
Slight ligament injury in knee
Slight ligament injury in hip
Slight ligament injury in back
Slight ligament injury in neck
Slight ligament injury in shoulder
Slight ligament injury in elbow
Slight ligament injury in wrist
Moderate ligament injury in ankle
Moderate ligament injury in knee
Moderate ligament injury in hip
Moderate ligament injury in back
Moderate ligament injury in neck
Moderate ligament injury in shoulder
Moderate ligament injury in elbow
Moderate ligament injury in wrist
Serious ligament injury in ankle
Serious ligament injury in knee
Serious ligament injury in hip
Serious ligament injury in back
Serious ligament injury in neck
Serious ligament injury in shoulder
Serious ligament injury in elbow
Serious ligament injury in wrist

Mild stroke
Moderate stroke
Serious stroke
Fatal stroke

***Infection**:*
temporary/occasional mild, of urinogenital tract
temporary/occasional mild, of eye
temporary/occasional mild, of ear
temporary/occasional mild, of lung
temporary/occasional mild, of skin
temporary/occasional serious, of urinogenital tract
temporary/occasional serious, of eye
temporary/occasional serious, of ear
temporary/occasional serious, of lung
temporary/occasional serious, of skin
chronic/regular mild, of urinogenital tract
chronic/regular mild, of eye
chronic/regular mild, of ear
chronic/regular mild, of lung
chronic/regular mild, of skin
chronic/regular serious, of urinogenital tract
chronic/regular serious, of eye
chronic/regular serious, of ear
chronic/regular serious, of lung
chronic/regular serious, of skin

Strained muscle/tendon in foot
Strained muscle/tendon in leg
Strained abdominal muscle
Strained muscle/tendon in arm
Strained muscle/tendon in hand
Strained muscle/tendon in neck

Strained muscle/tendon in jaw

Attempted suicide with full recovery
Attempted suicide with partial recovery
Suicide

Temporary immediate slight stress
Temporary immediate moderate stress
Temporary immediate severe stress
Chronic immediate slight stress
Chronic immediate moderate stress
Chronic immediate severe stress
Temporary delayed slight stress
Temporary delayed moderate stress
Temporary delayed severe stress
Chronic delayed slight stress
Chronic delayed moderate stress
Chronic delayed severe stress

Immediate slight tinnitus
Immediate moderate tinnitus
Immediate severe tinnitus
Delayed slight tinnitus
Delayed moderate tinnitus
Delayed severe tinnitus

Occasional slight unhappiness
Occasional moderate unhappiness
Occasional severe unhappiness
Regular slight unhappiness
Regular moderate unhappiness
Regular severe unhappiness

Occasional slight vertigo
Occasional moderate vertigo
Occasional severe vertigo
Regular slight vertigo
Regular moderate vertigo
Regular severe vertigo

APPENDIX C.
DERIVATION OF RATINGS OF SEVERITY OF HARM

C1. INTRODUCTION

The aim of this study was to develop a scale for measuring severity of harm which may be used in the evaluation of health and safety risks in buildings. The studies described here deal solely with how harms may be best classified and severity best measured, and do not cover related issues, for example the hazard/harm relationship, or the numbers of people who could be affected.

Measurement of harm can take several different forms, depending on the type of harm being examined, for example general well-being, disease and injury. Of these, it has been argued that an ordinal scale of injury is by far the easiest harm classification to understand and apply (Soby & Ball 1991). An example is provided by the Abbreviated Injury Scale (see Yates 1990) but this scale purposely makes no attempt to cover broader health outcomes, other than physical injury.

There are several approaches to the measurement of health states. Teeling Smith (1985) has identified three broad techniques: scales which give an overall score indicating severity/level of disease, profiles which provide measurements on several different states, and indices which combine measurements on different scales to produce a well-being score.

Several scales have been developed to measure a range of diseases and complaints, for example incapacity and depression scales. Many scales are, however, disease- and/or complaint-specific, for example psychological well-being or pain perception (see McDowell & Newell 1986) and as such cannot discriminate among or represent the variety of harms and injuries sustained in buildings.

A more general approach is offered by health profiles. These tend to give quantitative measurements of a number of distinct variables such as pain, mobility etc. The Nottingham Health Profile (Hunt et al 1980), for example, gives patient profiles on six areas: Physical mobility, Pain, Sleep, Energy, Social isolation, and Emotional reactions. This approach, as its name suggests, produces only a profile of the individual's state and the scores are not added in any way to give an overall score for the disease, injury, etc. which may produce these resultant states (Teeling Smith 1985).

A more complex approach to the measurement of well-being has been the development of health indices. These indices combine scores obtained from measurements of disability and discomfort into a global figure indicating current health state. While this is a more theoretically sound approach to measuring well-being of individual patients (Teeling Smith 1985) it may not be appropriate in developing a scale to measure all of the possible harms which may befall building users. Soby & Ball (1991) argue, for example, that clinically developed disease-specific indices are likely to be incompatible with aggregated statistical data, mainly due to person and situational specific dimensions (see Torrance et al 1982 for an example of such an index).

Existing measures deal only with specific injuries and/or diseases, while health profiles measure the effects of disease, injury, etc., without attempting to quantify the harm in question. These approaches are clearly not sufficient to describe all of the possible harms that may befall building users, but a more universal scale, appropriate to building use might be developed using similar techniques.

C2. SUMMARY OF APPROACH

Two approaches to the development of a measure of severity of representative harms (as derived in Appendix A), both appropriate in the evaluation of health and safety risks, are explored in the research reported here.

The first approach involves the direct comparison and rating of harms to construct an ordinal scale of severity. Essentially the harms are rated against a scale representing the severity of their impact. The scale points can be anchored using markers of severity, such as ‘mild, moderate, etc,' (OECD 1979). This method assumes that the concept of harm is unidimensional and that representative values of severity of harm can be identified and attributed with confidence. Scale values can be formed from mean scores given by a number of assessors for each harm.

The second approach involves identification of the different consequences of harm and rating types of harm against scales developed from those consequences. This is similar to the ‘health profile' approach in that it deals with harm outcomes. Various consequences have been identified in previous research; Fletcher et al (1988), for example, refer to the five Ds of negative health outcomes: death, disease, disability, discomfort and dissatisfaction

('destitution' is cited as a possible sixth outcome). In this approach subjects rate harms against each consequence scale. The data are then factor analysed and the results used as a statistical basis for combining scores.

The first approach may be defended if (a) the different consequences of harm, identified in the second approach, can be combined to form a general, or a single second order, factor, and (b) the empirical data and/or expert and public perceptions record more or less the same order of categories or types of harm against the individual consequences and/or second order factor.

The following sections describe the use of these two parallel approaches in scale development, provide a comparison of their results, and discuss the implications for the development of a general scale of harm.

C3. DIRECT RATINGS OF SEVERITY

C3.1 Introduction

The first approach to developing a measure of severity of harm was based on direct rating of harms, using a card sorting technique. It was carried out in three phases. First, a pilot study was carried out to test and refine the procedures to be used in the main study. This consisted of a sort of a sample of representative harms, to explore the criteria used by subjects when sorting the harms and to provide some preliminary data on the distribution of those harms. The main card sort was then carried out and each harm assigned a measure of severity based on its ratings. Finally, a reliability test was carried out. Each of these phases is described below.

C3.2 Pilot study

C3.2.1 Method

12 subjects took part in the pilot exercise. Subjects were drawn from the following three categories: (a) expert practitioners in the field of health and safety, (b) those with knowledge of, and/or interests in, health and safety practice and (c) those with medical knowledge. The subjects were derived from professional contacts with BRE or Loughborough University, including professionals attending short courses at the University. Comparison of the results with an independent sample, and with members of the public with no professional involvement in health and safety, is in hand.

A list of 293 harms, identified as 'representative harms' was used (see Appendix B). Each harm was printed on card (A6 size) and the cards were numbered on the reverse. Each scale number and marker were clearly printed on an A4 tent card. These cards were placed in ascending order from left to right, in front of the subject with an additional 'uncertain' card placed after the final scale card.

Subjects were given an instruction sheet, giving them the following instructions.

> *You will be presented with a selection of around 300 harms. These items represent likely harms to building users from both traumatic incidents (e.g. fire, falls, etc.) and long term exposure (e.g. building materials, noise, etc.).*
>
> *The items will be presented randomly in five groups. Your task is to sort the selections according to the following seven point scale:*
>
> *1 - Extremely Slight*
> *2 - Slight*
> *3 - Moderately Slight*
> *4 - Moderate*
> *5 - Moderately Severe*
> *6 - Severe*
> *7 - Extremely Severe*
>
> *The first 45 items will be used as a practice and then 62 items will be presented in each of the remaining groups. If you are uncertain of where a particular item fits the scale then it may be placed in the 'uncertain' pile. Please consider each harm individually and in terms of what is presented on the card.*

It was also explained to the subject that during the first, 'practice' sort (s)he should say why (s)he was placing cards in particular piles. It was not necessary for the subject to explain the reasoning behind every single decision, but to describe the basic strategies used. The strategies were noted and confirmed with the subject after the practice sort. After the first group of harms had been sorted, subjects were asked to try to place any items in the 'uncertain' pile. The remaining 248 harms were then presented in four blocks of 62 items, again randomly selected, with the 'uncertain' harms again placed if possible after each sort. The time for each sort was recorded, as well as the overall time to complete the task.

C3.2.2 Results

The mean completion time for the whole task was 59 minutes (range 46 to 75). The practice sort

took, on average, 17 minutes (range 14 to 21), and each of the remaining sorts took an average of 8 minutes (range 7 to 10).

All 12 subjects confirmed that they assessed each harm against its most likely consequences when rating its severity rather than just making a quick global decision. Specifically they commented on how much effect a particular harm would have on their everyday lives, how long it would take to recover, if recovery were possible, and/or the effects on the quality of life/standard of living. Three subjects also judged the severity of some of the harms in terms of their possible causes, and in terms of the implications of recurrence.

The different harms were tabulated in ascending order of severity with mean ratings, ranges and variances. Consideration of this distribution allowed the number of harms in the study to be reduced in two ways.

(a) The list was reduced by 18 by combining harms with similar meanings and similar means. If there were two similar harms with identical scores, ranges and variances, one was removed or the two were combined into one. Table C1 shows the new, composite, harms created in this way.

(b) A further 20 harms were removed where there was already a high degree of agreement among the subjects (these harms are listed in Appendix D). At the time of this work, the priority was to identify harms that would be markers for the scale points, hence the harms that were furthest from a scale point were preferentially removed.

These two actions resulted in a reduced list of 255 harms for the main study.

Subjects also made comments on the various aspects of the task; the comments can be summarised as follows.

- Most subjects, although told that around 300 harms were involved, thought that the whole task took longer than anticipated, and three thought that they had finished after the third sort.
- Nine subjects did not know what ‘grand mal' and ‘petit mal' seizures referred to.
- Eight subjects commented that, since they had been presented with the harms in five successive groups, they could not be sure of where they had placed similar harms in previous sorts and this made the task more difficult.
- Three subjects said that it was difficult not to place the harm in the category with a similar scale anchor to the harm's modifying adjective, for example ‘Regular Slight Dermatitis' may seem at first glance to go into the ‘Slight' category. However, subjects did endeavour to judge all harms according to their strategies and not simply match adjectives to anchor scales.
- Five subjects noted that they were more confident in placing harms of which they had some personal experience.

C3.3 Main card sort

C3.3.1 Method

Twenty subjects took part in the main card sort of the reduced list of harms. These subjects came from the same three groups as the pilot study subjects, namely (a) health and safety practitioners, (b) those with knowledge of, and/or interests in, health and safety practice and (c) those with medical knowledge.

The list of 255 harms derived from the pilot study was used. Each harm was printed on card (A7 rather than A6 size, to allow more cards to be placed side by side) and the cards were numbered on the reverse. The most problematic individual items in the pilot study were those referring to ‘grand mal' and ‘petit mal' seizures. These were clarified by inserting the word epileptic to give `petit mal epileptic seizure' etc.

The original seven point scale was made clearer by replacing ‘moderately severe' and `moderately slight' with ‘moderate/severe' and ‘moderate/slight' respectively. This better indicated that they represent the mid-points between moderate and the two other anchor points. The scale markers and numbers were again printed on A4 tent cards and placed in ascending order in front of the subject, with the ‘uncertain' marker placed after the final tent card.

Overall the procedure for the main card sort remained similar to that of the pilot study. It differed specifically in the following ways.

- The instruction to consider only what was presented on the card, in terms of overall or global impact on the individual, was further emphasised. The use of the possible consequences of the harms had been confirmed as a universal strategy in the pilot study, and this was reaffirmed in the instructions as the strategy which should be employed by those taking part.
- The subjects were given a list of all 255 harms to study for at least five minutes

before the start of the card sort. This was designed to give them a more realistic expectation of the length of the task and allow them to know what to expect in terms of the range of possible harms. Subjects were also allowed to choose two cards from each point on the scale, after the first sort, which they felt to be representative of that point. This was a further way of providing continuity between groups of items. These 'markers' were then retained for reference during the subsequent sorts.

- The preliminary distribution of harms allowed harms to be divided into seven groups of items with similar means. This, in turn, was used when compiling the blocks of items for presentation to subjects, ensuring a selection representative of the overall distribution in each group. The harms with the lowest variances were included in the first group of harms to allow subjects an 'easier' introduction to the task.

C3.3.2 Results

Each harm is shown in ascending order of severity in Appendix D with its mean rating, range and variance. This distribution of 255 harms is based on the combined responses from all 32 subjects taking part in both the pilot study and the main card sort. The list is divided into 12 clusters representing increments of 0.5 intervals on the severity scale. The clusters below [4] on the scale are, perhaps not surprisingly, characterised by the adjectives 'occasional' and 'temporary', while those from [5] and above are characterised by the adjectives 'immediate', 'regular', 'constant' and 'chronic'. The 'extremely severe' [7] point is made up of the only items with 100% agreement across all subjects and includes mostly fatal outcomes.

The final distribution of ratings is skewed toward the severe end of the scale, with over 60% of the harms rated above the mid-point of the scale.

It is possible to compare rated harms from different broad outcome categories. Table 2 in Chapter 2 shows seven different outcome categories and the harms relating to them. This table has been constructed using: (a) those items which clearly mark the ends of the scale, and (b) those items which clearly mark the centre of each of the original bands (1.5, 2.5, etc). The criteria of inclusion were: (a) means within 0.1 of the centre point, (b) associated variances as low as possible and certainly less than 1, and (c) ranges as low as possible and certainly 3 or less.

Although this table does allow direct comparison of different types of harm, none of the individual scales are in themselves complete. Comparison across categories at a particular level of severity is not, therefore, always possible.

An alternative approach would be to identify those harms closely associated with each scale point and use these as scale markers in the description of the final scale. This was accomplished by the development of an equal-appearing interval scale (Thurstone 1929).

The items with means closest to scale points and the smallest semi-interquartile (S-I) ranges can be selected as scale markers. Table C2 shows the equal-appearing interval scale, which can be developed from the card sort. The semi-interquartile ranges should be as small as possible and only items with S-I ranges less than or equal to 0.5 have been included (Dane 1990).

C3.4 Reliability test

C3.4.1 Method

Thirty subjects, drawn from the same populations as in the pilot and main studies, were involved in the reliability test. Seven items (see Table C3) from the equal-appearing interval scale were used to test reliability of the scale in a questionnaire format. Items were rated on the same seven point scale used in the main study, ranging from 'extremely slight' [1] to 'extremely severe' [7].

Subjects were asked to rate each of the scale markers on two occasions, separated by at least 24 hours. No indication was given to subjects on the first occasion that they would be asked to complete it once more.

C3.4.2 Results

Table C3 shows the means for the scale markers from the main sort and both stages, pre and post, of the reliability test. The test-retest correlation coefficients (Pearson's r) are also shown. The rank order of all the harms remained the same and all reliability coefficients shown are of an acceptable level. The highest coefficients are at the two extremes of the scale.

C3.5 Conclusions on direct ratings of severity

The first approach to developing a measure of severity of harm provided a simple scale. The distribution of harms was, however, skewed toward the more severe end of the scale. This may have been expected since, by their very nature, harms are not desirable and therefore there may be some resistance to describing them as slight.

There was also relatively little agreement between subjects on some items, and this may have been because of a shortage of information. Subjects therefore made judgements of these items dependent on their own experience and/or imagination. For example *'Minor puncture/cuts to hands'* (variance=1.10; range=5) could be visualised very differently depending on the subject's experience and imagination.

The main criticism of this scale is that, in the form presented in the reliability test, it is composed of a mixture of types of harm. While the study aimed to provide comparisons across types of harms, examples are not given of each of the different types of harm from across the full range of the scale. Separating out categories of harm, as in Table 2, presents scales that are more easily appreciated, but each of these is incomplete and could not, therefore, be used on its own.

The scale developed here may, since it was designed to encompass all types of harm, reflect judgements based on different, discrete dimensions of severity. If different categories of harm were judged on discrete criteria then an overall comparison of all harms would be problematic. This approach can however be defended if, as already stated, different consequences of harm can be combined to form a single factor or property, and if the resulting scale records more or less the same order of harm. The following section describes the development of such a scale.

C4. SCALING BY CONSEQUENCE

C4.1 Introduction

The second approach to the development of a measure of severity of harm, based on rating harms against consequences was carried out in two phases. First, consequence constructs were identified using 'triadic elicitation' (see below). Then a selection of harms was rated using consequence scales developed from these constructs. The results of this exercise were analysed and a scale based on consequence ratings produced.

C4.2 Generation of constructs

C4.2.1 Method

30 subjects took part in this exercise. These were drawn from the same populations as in the previous studies with the addition of Health Psychology and Ergonomics students who had an interest in health and safety.

The harms in Appendix B were stripped of all modifying adjectives, for example 'slight', 'regular', etc, to produce a list of 126 basic harms. These were printed on cards (A7 size) and the cards were numbered on the reverse. These harms were split randomly into two groups (A and B) of 63 harms each, for presentation to subjects.

Subjects were told that they would be presented with 63 harms in triads (groups of three), and that these represented harms which may befall an individual. One triad was chosen and subjects were asked to say in what way two of the three harms were similar, but differed from the third, in terms of their consequences. The description of similarity and difference represents a 'construct'. An example was shown (from the other group of 63 harms) to ensure subjects understood what was expected. This method of triads is similar to the preliminary stages of repertory grid analysis (e.g. see Honey 1979).

Triads of harms were then chosen at random and presented to the subject. Constructs were noted down alongside the harms presented. This was repeated until all the harms in the chosen group had been dealt with in the same way. Twelve subjects were tested on each group, A and B. The remaining 6 subjects were tested on random mixtures of 63 items taken from both groups. This resulted in a total of 630 groupings of three harms being presented to subjects.

C4.2.2 Results

Table C4 shows the results of the triadic elicitation exercise based on ratings of 630 triad groupings. The eight consequence constructs listed represent those identified by the sample of subjects more than 40 times, and at least once by each subject.

A further 10 subjects were presented with these constructs and asked to group together any that could have similar meanings. The construct 'Has a long term impact' was placed with 'Restricts mobility permanently' four times and 'Results in death' six times. 'Incapacitates totally' was identified with 'Restricts mobility permanently' by nine subjects. 'Changes behaviour' was identified with 'Changes lifestyle' seven times. Eliminating the grouped constructs with the lower frequencies leaves five main constructs (marked * in Table C4) to be developed into scales.

The consequences identified here and judged as distinct resemble some of the outcomes identified in previous research (e.g. Fletcher et al 1988), as well as the dimensions involved in health profiling (e.g. Hunt et al 1980).

C4.3 Rating by consequence

C4.3.1 Introduction

The consequence constructs which were found to underpin harms may represent distinct factors and subjects may, therefore, be using these factors in different ways in forming their judgements; calling into question the utility of a universal relative scale of harms. Scaling a selection of harms using consequence scales will allow investigation of the factors that underpin a subject's decisions. If a unidimensional scale can then be derived from the individual consequence scales used then subjects may be basing their judgements on one factor, endorsing the development of a single scale.

C4.3.2 Method

There were 138 subjects, 5 of whom were also involved in a pilot study. Subjects were drawn from the same groups as in the previous studies.

The procedure was piloted using four harms chosen at random from the list shown in Appendix D *(Simple fracture of legs, Severe birth defect, Slight concussion* and *Moderate bruising to legs).* These harms were placed at the end of the sentence "How true are each of the following possible consequences for someone who has suffered [description of harm]". This sentence was followed by the five consequence constructs identified in the triad exercise, each paired with a seven point scale. The scale was designed to measure how applicable each negative outcome was to the harm in question; it was a semantic differential scale ranging from [1] 'Definitely not true' to [7] 'Definitely true'.

Five subjects completed the pilot and all felt that a 'Don't know' or 'Can't say' point should be included. This point was added after the final point of each scale.

The 13 harms shown in Table C5 were chosen from the list in Appendix D for inclusion in the final questionnaire. One harm was taken from each scale point and each mid-point, representing, as much as possible, the broad outcome categories set out above (see Table 2). These harms were used in the same format as in the pilot study.

The resulting data were factor analysed using principal components analysis (see Ferguson & Cox 1993). The five consequence items were analysed regardless of their related harm to explore the possibility of deriving either (a) a general factor covering all of the consequence scales, or (b) a single second order factor summarising any primary factor structure.

C4.3.3 Results

In order for the data set to be appropriate for factor analysis it must meet certain requirements. First, the minimum number of subjects recommended for such an analysis is between 100 and 200 (Ferguson & Cox 1993). The sample size in this case, 138, met that minimum requirement. The next stage of the analysis involves testing the correlation matrix derived from the data. Two tests can be applied to this end (Dzuiuban & Shirkey 1974): the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett test of sphericity (BS). The KMO indicates whether the relationships within the matrix can be accounted for by a smaller number of factors and a minimum value of 0.5 is required. The BS indicates whether there are significant discoverable relationships in the data. The KMO value for this data set was 0.87 and the BS was significant at the 0.001 level, showing that the matrix is appropriate for factor analysis.

Factor analysis of the five consequence items produced one primary factor. One heuristic suggested for determining the number of factors to extract is the Kaiser 1 rule (see Ferguson & Cox 1993), which suggests the extraction of as many factors as there are eigenvalues greater than one. Only one eigenvalue resulting from this analysis was greater than one (3.84), supporting the extraction of one factor. This factor accounted for 76.9% of the overall variance. The fact that one factor was found to underpin all consequence constructs endorses the universal card sort methodology; if harm is unidimensional then all harms can be sorted into a relative scale. Further validation, using more subjects and a longer list of harms, would be desirable but was not possible within the scope of this study.

The acceptable level for the internal reliability of a scale (Cronbach's alpha) is 0.7 and above. The alpha for the five consequence items combined, irrespective of the harm to which they relate, is 0.91.

Adding the scores from each consequence scale provides an overall scale ranging from 5 representing the absence of negative consequences, to 35 representing the presence of all negative consequences. Measuring severity in terms of consequences in this way allows a severity value to be calculated for each of the harms. Individual alphas can also be calculated to indicate the reliability of the judgements for each harm. Table C5 shows the harms with their mean values, ranges and alpha values, indicating the degree of consensus.

Seven of the alphas reported for the consequence scales are below the acceptable level. This indicates that subjects showed less agreement on the items. Subjects may have found it more difficult to rate items if they considered some consequences more important than others, depending on the harm in question.

C4.4 Construct severity scale development

Dividing the scales arrived at in the questionnaire study by the number of constructs (5) would provide seven point scale values comparable with those derived from the card sort. This assumes, however, that each consequence construct is equally important when judging severity. This assumption was tested by asking 15 subjects to rate how important each consequence construct was when judging severity of personal harm, using a five point scale, where [1] represented 'very unimportant' and [5] represented 'very important'. The results of this exercise provide the mean construct weightings shown in descending order in Table C6.

Multiplying each construct score by its weighting and then adding these weighted scores gives an overall severity score for each harm. These score range from 20.3 to 142.1 and can be standardised to represent a seven point scale. This standardised scale was then compared to the scale derived from the simple addition of consequence scales described above.

The correlation between the two scales was $r = 0.96$, which was significant at the 0.01 level. This large coefficient indicated that there was no practical difference between the two methods of scale standardisation in terms of outcome scores. The weighting method was used to calculate scores for comparison with the direct rating scale.

C4.5 Conclusions on scaling by consequence

The constructs judged as distinct in the triadic elicitation exercise contribute to a single 'severity of harm' factor and are similar to the negative outcomes identified by other research. The constructs elicited directly and indirectly reflect some of the 'D's of negative health outcome referred to by Fletcher et al (1988) and mentioned above. Death is represented directly, while disability is reflected, to some degree, by 'Restricts mobility permanently' and 'Affects physical performance'. Similarly two of the dimensions form the main part of the Nottingham Health Profile, 'Social isolation' and 'Physical mobility' are reflected in the consequence constructs identified.

C5. SCALE COMPARISON

C5.1 Introduction

This section deals with comparison of the scales derived from the two approaches. As stated earlier, the second approach to developing a scale for measuring severity of harm tests the foundations of the first approach. The consequences identified can be described by a single primary factor, but in order for the relative scale to be supported fully the two studies must produce more or less the same order of categories or types of harm.

C5.2 Procedure

Table C7 shows the thirteen harms involved in both the card sort and the questionnaire study, with their scale values and relative positions from each. Comparing scale values from the two studies shows that all outcome scale values are within one scale point of their corresponding card sort values. The markers at the ends of the scales match very closely.

The scales used were not identical and relative ranks may provide a better basis for comparison. Those items in the lower and mid parts of the scale have indeed changed their order but only slightly, with *'Short term slight high blood pressure'* the only item to move two places.

There is, however a high level of agreement between the scales ($r=0.96$). With such a high correlation, any difference between the scales could easily result from imprecision in the study, rather than a fundamental difference between the two scales.

C5.3 Conclusions on scale comparison

The comparison of the two approaches shows that there is very good agreement at both ends of the scales and an acceptable degree of agreement on the harms in the centre. There is a strong linear relationship between the two scales, suggesting that the results of the card sort exercise may be relied on to produce a general scale of severity. The comparison of scales indicates a need for wider bands than those presented, for example, in the equal-appearing interval scale.

C6. CONCLUSIONS

In order to develop a scale of severity appropriate to building use a wide range of harms which might befall building users must be considered. Direct

ratings of severity provide a relatively straightforward scale for measuring severity but the results of the scaling by consequence exercise, described above, must be taken into account in formation of a scale.

The scale shown in Table 3 in Chapter 2 provides one solution to the formation of a scale of severity of harm. This scale covers all types of harm, is derived from the direct ratings of those harms, and is based on the original seven point scale of severity. This scale also has the following characteristics:

- there is reasonable similarity of the severity of harms within bands;
- the bands can easily be named (since they correspond to the top four scale points, plus a lowest band);
- the most severe band clearly represents death and comparable levels of harm.

The lowest band is however composed largely of harms for which statistics would not be kept, and which are at a level unlikely to affect national legislation. The recommendation is therefore that the top four bands of the five-band scale be adopted for the purpose of comparing harms that occur, and hazards that exist, within buildings. The full five-band scale could be used for more local evaluations, for example involving specific buildings or regions, or in assessments related to well-being and productivity. The scale could be further validated against the judgements of non-experts, using only the more widely known and understood harms, and could be rendered more easily applicable by quantifying the differences in severity between bands.

REFERENCES

Dane, F C (1990) *Research Methods.* Brooks/Cole Publishing: Pacific Grove, Calif.

Dzuiuban, C & Shirkey, E (1974) When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, **81,** 358-361.

Ferguson, E & Cox, T (1993) Exploratory factor analysis: A user's guide. *International Journal of Selection and Assessment*, **1**, 84-93.

Fletcher, R H, Fletcher, S W & Wagner, E H (1988) *Clinical Epidemiology.* Williams and Williams: Baltimore

Honey, P (1979) The repertory grid in action. *Industrial and Commercial Training*, **11(9)**, 358-369

Hunt, S M, McKenna, S P, McEwen, J, Backett, E M, Williams, J & Papp, E (1980) A quantitative approach to perceived health status: A validation study. *Journal of Epidemiology and Community Health*, **34,** 281-286

McDowell, I & Newell, C, (1986) *Measuring Health: A Guide to Rating Scales and Questionnaires.* Oxford University Press: New York.

OECD, (1979) *Severity Weighting of Data on Accidents Involving Consumer Products*. OECD, Paris.

Soby, B A & Ball D J (1991) *Consumer Safety and the Valuation of Life and Injury*. Consumer Safety Unit, Department of Trade and Industry, London.

Teeling Smith, G (1985) *Health Measurement.* Office of Health Economics: London.

Torrance, G W, Boyle, M H & Horwood, S P (1982) Application of multi-attribute utility theory to measure social preferences for health states. *Operations Research,* **30,** 1043-1069.

Thurstone, L L (1929) Theory of attitude measurement. *Psychological Review*, **36**, 222-241.

Yates, D W (1990) Scoring systems for trauma. *British Journal of Medicine*, **301**, 1090-1094.

Table C1. New harms created from merger of harms used in pilot study

HARM (NEW HARM IN ITALIC)	MEAN	RANGE	VAR
Occasional slight annoyance/dissatisfaction			
10. Occasional slight annoyance	1.33	1 - 3	0.42
98. Occasional slight dissatisfaction	1.33	1 - 3	0.42
Constant serious annoyance/dissatisfaction			
12. Constant serious annoyance	4.42	3 - 6	0.99
99. Regular serious dissatisfaction	4.42	3 - 6	0.63
Asphyxiation resulting in major brain damage/death			
17. Asphyxiation with major brain damage	7.00	7	0.00
18. Asphyxiation resulting in death	7.00	7	0.00
Loss of arms			
23. Loss of arms in accident	6.67	6 - 7	0.24
24. Loss of arms because of burns	6.83	6 - 7	0.15
Loss of arms and legs			
25. Loss of arms and legs by amputation because of injury	6.92	6 - 7	0.08
26. Loss of arms and legs by amputation because of disease	7.00	7	0.00
Loss of finger			
29. Loss of finger by amputation because of disease	5.50	4 - 7	1.00
30. Loss of finger because of burns	5.42	3 - 7	1.36
Loss of foot			
32. Loss of foot by amputation because of injury	6.42	5 - 7	0.63
33. Loss of foot by amputation because of disease	6.42	5 - 7	0.45
Superficial blisters other than on face			
49. Superficial blisters on hands	2.58	2 - 4	0.45
50. Superficial blisters other than on face or hands	2.58	2 - 4	0.45
Serious blisters on face or hands			
51. Serious blisters on face	4.08	3 - 5	0.63
52. Serious blisters on hands	4.08	3 - 5	0.81
Gangrene in hand/toe			
126. Gangrene in hand	5.83	4 - 7	0.70
127. Gangrene in toe	5.42	3 - 7	1.72
Occasional slight irritation of nose/rhinitis			
184. Temporary/occasional slight irritation of nose	1.83	1 - 3	0.33
217. Occasional slight rhinitis	1.91	1 - 3	0.27
Regular severe irritation of nose/rhinitis			
186. Chronic/regular severe irritation of nose	4.58	4 - 6	0.63
219. Regular severe rhinitis	4.55	2 - 6	3.72
Permanent paralysis below neck			
215. Permanent paralysis below waist	6.67	6 - 7	0.24
216. Permanent paralysis below neck	6.92	6 - 7	0.08
Strained muscle/tendon in foot/leg			
277. Strained muscle/tendon in foot	2.67	1 - 4	1.33
278. Strained muscle/tendon in leg	2.75	2 - 4	0.39
Fatal stroke/heart attack			
151. Immediate fatal heart attack	7.00	7	0.00
262. Fatal stroke	7.00	7	0.00
Occasional slight irrational fear (e.g. of spiders, enclosed spaces)			
53. Mild claustrophobia	2.25	1 - 3	0.39
119. Occasional slight irrational fear (e.g. of spiders, enclosed spaces)	2.25	1 - 4	0.75
Loss of dominant hand			
28. Loss of dominant hand because of burns	6.08	4 - 7	0.63
27. Loss of dominant hand in accident	6.58	6 - 7	0.27
Chronic Severe Wheezing/Breathlessness (Immediately Induced)			
20. Onset of serious asthmatic condition	5.75	5 - 7	0.57
73. Chronic severe wheezing/breathlessness (immediately induced)	5.75	4 - 6	0.39

Table C2. Equal-appearing interval scale

SCALE POINT	ITEM VALUE	S-I RANGE	ITEM
1	1.06	0.0	Slight graze to hand
2	2.03	0.5	Occasional slight irritation of the nose
3	2.97	0.5	Regular slight fever
	3.00	0.0	Regular slight irritation of the nose
4	4.00	0.0	Short term amnesia
5	5.00	0.0	Major hernia
6	6.00	0.5	Severe hypothermia
7	7.00	0.0	Asphyxiation resulting in major brain damage/death
	7.00	0.0	Leukaemia - non-treatable
	7.00	0.0	Fatal immediate brain damage
	7.00	0.0	Acute fatal bronchitis
	7.00	0.0	Fatal electric shock
	7.00	0.0	Permanent loss of consciousness
	7.00	0.0	Fatal hyperthermia
	7.00	0.0	Fatal internal abdominal damage
	7.00	0.0	Fatal stroke/heart attack
	7.00	0.0	Suicide

Table C3. Scale marker means

ITEM	SORT MEAN	PRE-TEST	POST-TEST	CORRELATION COEFFICIENT (R)
1 Slight graze to hand	1.06	1.07	1.07	1.000**
2 Occasional slight irritation of nose	2.03	1.90	1.93	0.814**
3 Regular slight fever	2.97	3.20	3.23	0.754**
4 Short term amnesia	4.00	4.00	4.13	0.845**
5 Major hernia	5.00	5.27	5.27	0.696**
6 Severe hypothermia	6.00	6.13	6.17	0.933**
7 Permanent loss of consciousness	7.00	7.00	7.00	1.000**

** $p < .001$

Table C4. Triadic consequence constructs

CONSTRUCT		FREQUENCY
1	Results in death*	81
2	Changes lifestyle*	73
3	Restricts mobility permanently*	69
4	Has a long-term impact	67
5	Incapacitates totally	58
6	Affects physical performance (in and out of work)*	56
7	Inhibits social interaction*	50
8	Changes behaviour	45

Table C5. Harms in ascending order from the questionnaire study

HARM	MEAN	RANGE	ALPHA
Slight graze to hand	5.14	5-7	.70
Occasional slight irritation of the nose	8.82	5-15	.53
Occasional slight discomfort	10.44	5-21	.88
Short term slightly high blood pressure	14.19	5-25	.79
Temporary slight backache	14.57	8-28	.81
Regular slight fever	15.84	6-25	.63
Short term amnesia	19.59	11-30	.68
Immediate moderate tinnitus	19.70	8-33	.67
Major hernia	25.43	17-34	.45
Constant serious anxiety	25.76	15-32	.51
Severe hypothermia	26.13	16-35	.75
Loss of arms	29.93	25-35	.23
Permanent loss of consciousness	34.72	34-35	.86

Table C6. Construct weightings

CONSTRUCT (AS USED IN QUESTIONNAIRE)	WEIGHTING
Results in death	4.9
Results in permanent restriction on mobility	4.3
Restricts physical performance	4.1
Results in changes to lifestyle	3.7
Inhibits social interaction	3.3

Table C7. Comparison of scales

	CARD SORT		SCALING BY CONSEQUENCE	
ITEM	Scale Value	Rank	Scale Value	Rank
Slight graze to hand	1.06	1	1.03	1
Occasional slight discomfort	1.62	2	2.05	3
Occasional slight irritation of the nose	2.03	3	1.69	2
Temporary slight backache	2.47	4	2.88	5
Regular slight fever	2.97	5	3.16	6
Short term slight high blood pressure	3.47	6	2.85	4
Short term amnesia	4.00	7	3.83	8
Immediate moderate tinnitus	4.47	8	3.78	7
Major hernia	5.00	9	5.07	10
Constant serious anxiety	5.47	10	5.04	9
Severe hypothermia	6.00	11	5.32	11
Loss of arms	6.50	12	5.91	12
Permanent loss of consciousness	7.00	13	6.98	13

APPENDIX D.
LIST OF HARMS AND MEAN RATINGS IN ASCENDING ORDER[1]

HARM		MEAN	VAR.	RANGE
109.	Slight graze to hand	1.06	0.06	1.00
108.	Slight graze to arm/leg	1.22	0.24	2.00
174.	Occasional slight loss of appetite	1.31	0.35	2.00
93.	Occasional slight fatigue	1.59	0.57	3.00
78.	Occasional slight discomfort	1.62	0.44	2.00
9.	Occasional slight annoyance/dissatisfaction	1.66	0.62	2.00
36.	Occasional insect bites	1.78	0.56	3.00
163.	Temporary/occasional slight irritation of skin	1.81	0.80	3.00
111.	Occasional slight headache	1.91	0.67	2.00
251.	Occasional slight unhappiness	1.94	0.64	2.00
157.	Occasional slight irritation of nose/rhinitis	2.03	0.48	2.00
167.	Occasional mild cough/cold	2.03	0.74	3.00
99.	Occasional slight irrational fear (e.g. of spiders, enclosed spaces)	2.09	0.54	3.00
200.	Minor puncture/cuts to hands	2.16	1.10	5.00
171.	Occasional slight lack of privacy	2.19	0.74	3.00
202.	Minor puncture/cuts to leg	2.25	0.65	4.00
234.	Temporary/occasional mild infection of ear	2.25	1.03	3.00
67.	Minor dislocated finger	2.28	1.37	3.00
146.	Occasional slight impairment of concentration	2.28	0.47	3.00
176.	Slightly low birth weight	2.28	1.18	3.00
12.	Occasional slight anxiety	2.31	1.32	3.00
214.	Temporary mild sleep disorder	2.34	0.68	3.00
87.	Occasional slight eye strain	2.37	0.95	4.00
110.	Bad graze to face	2.37	1.02	4.00
253.	Occasional slight vertigo	2.37	0.50	4.00
41.	Superficial blisters other than on face	2.38	0.50	3.00
235.	Temporary/occasional mild infection of skin	2.38	1.02	4.00
101.	Occasional slight fever	2.41	0.64	3.00
178.	Occasional slight menstrual disorder	2.42	0.98	4.00
187.	Temporary/occasional slight irritation of throat*	2.42	0.27	1.00
46.	Moderate bruising to legs	2.44	1.03	4.00
18.	Temporary slight backache	2.47	0.58	3.00
143.	Occasional slight impairment of children's play	2.50	1.16	3.00
217.	Occasional mild sleep deprivation	2.50	1.16	4.00
151.	Occasional slight restriction of social behaviour	2.53	0.71	3.00
66.	Minor dislocated toe	2.53	1.16	4.00
81.	Occasional immediate slight drowsiness	2.59	1.15	4.00
3.	Occasional irritability/intolerance towards other occupants	2.63	0.76	3.00
94.	Regular slight fatigue	2.63	0.82	3.00
75.	Occasional slight dermatitis	2.66	0.94	3.00
154.	Temporary/occasional slight irritation of eyes	2.66	1.65	5.00
230.	Slight ligament injury in ankle	2.66	0.49	3.00
85.	Slight electric shock	2.75	1.94	5.00
141.	Occasional/temporary slight impairment of sexual behaviour	2.78	0.63	3.00
198.	Moderate puncture/cuts to ankles/feet	2.81	0.80	3.00
240.	Strained muscle/tendon in foot/leg	2.81	1.00	3.00
223.	Temporary slight speech impairment	2.84	1.17	4.00
246.	Temporary delayed slight stress	2.84	0.97	3.00
218.	Occasional moderate sleep deprivation	2.87	1.08	4.00
45.	Moderate bruising to face	2.88	1.08	4.00
182.	Temporary/occasional slight muscular weakness	2.88	1.53	4.00
50.	Minor burns/scalds to legs/ankles	2.94	1.09	5.00

[1]Items marked * were used only in the pilot study (all had variance < 0.3 and interquartile range ≤ 1).

	HARM	MEAN	VAR.	RANGE
102.	Regular slight fever	2.97	0.81	3.00
123.	Hairline fracture of nose	3.00	2.13	5.00
160.	Regular slight irritation of nose/rhinitis	3.00	0.40	2.00
139.	Occasional/temporary slight impairment of judgement	3.03	0.93	3.00
254.	Regular slight vertigo	3.03	0.93	3.00
58.	Temporary slight wheezing/breathlessness	3.06	1.61	5.00
122.	Hairline fracture of toes	3.12	1.02	3.00
96.	Occasional slight fear of real threat (e.g. crime, children's safety)	3.23	1.25	4.00
241.	Strained abdominal muscle	3.25	0.97	3.00
76.	Regular slight dermatitis	3.28	0.85	4.00
73.	Occasional moderate diarrhoea/vomiting	3.31	1.77	5.00
266.	Slight ligament injury in wrist*	3.33	0.24	1.00
10.	Regular moderate annoyance/dissatisfaction	3.37	0.95	4.00
158.	Occasional severe irritation of nose/rhinitis	3.37	1.02	4.00
147.	Regular moderate impairment of concentration	3.38	1.47	4.00
250.	Delayed slight tinnitus	3.41	1.67	4.00
208.	Slight temporarily restricted child development (social/ emotional/cognitive)	3.42	1.32	3.00
68.	Minor dislocated shoulder	3.44	0.90	3.00
70.	Subclinical short-term depression	3.44	0.51	3.00
206.	Slight birth defect	3.44	1.80	5.00
232.	Moderate ligament injury in elbow	3.44	0.77	3.00
63.	Mild temporary confusion/disorientation	3.47	1.10	3.00
127.	Short-term slight high blood pressure	3.47	0.90	3.00
180.	Occasional moderate migraine	3.47	0.90	5.00
112.	Regular moderate headache	3.50	1.61	5.00
191.	Temporary delayed slight personality change	3.50	1.42	5.00
135.	Immediate temporary impaired hearing in one ear	3.53	1.42	5.00
244.	Temporary immediate moderate stress	3.53	0.45	3.00
161.	Chronic/regular slight irritation of throat	3.56	1.48	5.00
88.	Regular slight eye strain	3.59	1.02	4.00
189.	Temporary immediate slight personality change	3.59	1.60	4.00
84.	Regular delayed moderate drowsiness	3.62	1.21	4.00
121.	Simple fracture of fingers	3.63	1.53	4.00
201.	Moderate puncture/cuts to head/face	3.66	0.75	4.00
13.	Regular moderate anxiety	3.69	1.06	4.00
204.	Moderate puncture/cuts to torso	3.69	1.00	5.00
24.	Loss of finger/thumb/toe nail by surgical removal	3.72	1.89	5.00
149.	Minor hernia	3.72	0.66	3.00
166.	Chronic/regular redness/rash other than on face or hands	3.72	0.72	2.00
224.	Chronic slight speech impairment	3.75	1.87	5.00
228.	Slight slipped disc	3.78	1.34	5.00
28.	Mild temporary behaviour disorder (immediately following hazardous event)	3.78	1.79	5.00
17.	Onset of mild asthmatic condition	3.84	0.65	2.00
196.	Temporary slight psychosis	3.84	1.81	5.00
172.	Regular moderate lack of privacy	3.87	0.92	4.00
56.	Slight concussion	3.88	0.89	3.00
26.	Temporary loss of hair because of disease	3.91	1.51	4.00
193.	Occasional mild pneumonia	3.91	0.99	4.00
231.	Moderate ligament injury in hip	3.91	0.54	3.00
164.	Chronic/regular moderate irritation of skin	3.94	1.03	4.00
152.	Regular moderate restriction of social behaviour	3.97	0.87	4.00
8.	Short term amnesia	4.00	0.71	4.00
48.	Severe bruising to torso	4.00	1.55	4.00
168.	Regular serious cough/cold	4.00	1.03	3.00
83.	Occasional delayed severe drowsiness	4.03	1.26	5.00

HARM		MEAN	VAR.	RANGE
124.	Simple fracture of jaw	4.06	1.61	4.00
79.	Occasional severe discomfort	4.09	0.93	3.00
54.	10% burns/scalds	4.13	1.85	4.00
100.	Regular severe irrational fear (e.g. of spiders, enclosed spaces)	4.13	2.45	6.00
113.	Simple fracture of ankles/feet	4.13	1.02	4.00
209.	Moderate temporarily restricted child development (social/ emotional/cognitive)	4.13	0.92	3.00
51.	Moderate burns/scalds to feet	4.16	1.10	4.00
155.	Chronic/regular slight irritation of eyes	4.16	1.17	4.00
31.	Benign tumour(s)	4.19	1.32	4.00
165.	Chronic/regular redness/rash on face	4.19	1.13	4.00
212.	Slight immunological sensitisation to chemical/biological agents	4.22	1.66	4.00
37.	Immediate temporary impaired vision in one eye	4.25	0.71	3.00
90.	Brief loss of consciousness	4.25	1.61	5.00
42.	Serious blisters on face or hands	4.28	0.85	3.00
131.	Mild hyperthermia	4.28	1.37	4.00
118.	Hairline fracture of pelvis	4.34	0.81	3.00
220.	Immediate temporary partial dependency	4.34	0.62	3.00
144.	Regular moderate impairment of children's play	4.38	1.02	3.00
185.	Temporary paralysis of foot/feet	4.41	0.96	4.00
175.	Regular severe loss of appetite	4.42	2.12	4.00
159.	Regular severe irritation of nose/rhinitis	4.44	0.83	4.00
236.	Temporary/occasional serious infection of eye	4.44	0.58	4.00
136.	Gradual temporary impaired hearing in both ears	4.47	0.58	3.00
173.	Regular severe lack of privacy	4.47	0.64	3.00
247.	Chronic delayed moderate stress	4.47	0.39	2.00
248.	Immediate moderate tinnitus	4.47	0.71	3.00
177.	Very low birth weight	4.50	1.48	5.00
199.	Serious puncture/cuts to arm/wrist	4.50	0.97	4.00
238.	Chronic/regular serious infection of urinogenital tract	4.50	1.23	4.00
53.	Moderate burns/scalds to torso	4.53	0.97	3.00
61.	Acute moderate bronchitis	4.53	1.16	4.00
179.	Regular severe menstrual disorder	4.55	0.72	3.00
11.	Constant serious annoyance/dissatisfaction	4.56	1.16	4.00
97.	Regular moderate fear of real threat (e.g. crime, children's safety)	4.58	0.98	4.00
116.	Compound, complicated or comminuted fracture of wrist/hands	4.62	0.89	3.00
203.	Serious puncture/cuts to neck	4.63	1.21	3.00
211.	Slight retinal haemorrhage	4.66	1.39	5.00
162.	Chronic/regular severe irritation of throat	4.66	0.75	3.00
192.	Chronic delayed moderate personality change	4.68	1.63	4.00
197.	Temporary moderate psychosis	4.68	1.03	4.00
138.	Gradual permanent deafness in one ear	4.69	2.03	5.00
49.	Severe multiple bruising	4.72	1.24	4.00
89.	Regular severe eye strain	4.72	1.18	4.00
107.	Stomach/duodenal ulcer	4.72	0.79	3.00
183.	Chronic/regular moderate muscular weakness	4.72	1.05	3.00
252.	Regular severe unhappiness	4.72	1.37	4.00
95.	Regular severe fatigue/exhaustion	4.75	1.03	3.00
85.	Serious dislocated jaw*	4.75	0.20	1.00
194.	Occasional moderate pneumonia	4.77	0.58	3.00
23.	Loss of finger	4.78	2.69	6.00
40.	Gradual temporary blindness in one eye	4.78	0.56	3.00
47.	Severe bruising to head/neck	4.78	1.27	4.00
64.	Mild chronic confusion/disorientation	4.78	0.76	3.00
215.	Chronic moderate sleep disorder	4.78	0.95	4.00
233.	Serious ligament injury in knee	4.78	0.50	3.00
237.	Chronic/regular mild infection of lung	4.78	1.08	3.00
71.	Severe short-term depression	4.81	0.87	4.00

HARM		MEAN	VAR.	RANGE
130.	Mild angina	4.81	0.87	5.00
4.	Occupant occasionally violent when in building	4.84	0.65	3.00
60.	Chronic moderate wheezing/breathlessness (gradually induced)	4.84	0.59	3.00
205.	Multiple serious puncture/cuts	4.84	2.33	4.00
104.	Occasional petit mal epileptic seizure	4.87	0.50	3.00
27.	Permanent loss of hair because of burns	4.88	1.60	4.00
69.	Serious dislocated knee	4.88	1.08	3.00
7.	Alienation from society (as an immediate result of occupying a certain home)	4.90	1.02	4.00
145.	Regular severe impairment of children's play	4.90	1.62	5.00
133.	Moderate hypothermia	4.94	1.03	3.00
128.	Chronic moderate high blood pressure	4.97	0.68	3.00
142.	Regular/chronic severe impairment of sexual behaviour	5.00	1.87	5.00
150.	Major hernia	5.00	0.26	2.00
153.	Regular severe restriction of social behaviour	5.00	0.71	3.00
15.	Asphyxiation with full recovery	5.03	1.19	4.00
29.	Moderate chronic behaviour disorder (developing over time)	5.03	0.61	4.00
80.	Regular severe discomfort	5.03	0.68	3.00
225.	Chronic severe speech impairment	5.06	1.48	5.00
38.	Gradual permanent impaired vision in one eye	5.09	0.67	3.00
52.	Severe burns/scalds to hands	5.16	1.10	3.00
77.	Regular severe dermatitis	5.16	0.72	4.00
148.	Regular severe impairment of concentration	5.16	0.46	2.00
156.	Chronic/regular severe irritation of eyes	5.19	0.74	3.00
5.	Regular aggressive behaviour when in building	5.19	0.54	2.00
43.	Severe claustrophobia	5.22	1.02	5.00
82.	Regular immediate severe drowsiness	5.22	0.63	4.00
98.	Regular severe fear of real threat (e.g. crime, children's safety)	5.23	1.25	4.00
249.	Immediate severe tinnitus	5.25	0.97	4.00
255.	Regular severe vertigo	5.28	0.92	3.00
181.	Regular severe migraine	5.31	0.67	3.00
120.	Simple fracture of skull	5.34	0.81	3.00
245.	Chronic immediate severe stress	5.34	0.49	3.00
169.	Moderate internal abdominal damage	5.38	0.37	3.00
219.	Regular serious sleep deprivation	5.38	0.89	4.00
186.	Temporary paralysis of arms	5.39	1.18	4.00
74.	Regular severe diarrhoea/vomiting	5.41	0.70	4.00
114.	Compound, complicated or comminuted fracture of arms/shoulder	5.41	0.70	4.00
184.	Chronic/regular severe muscular weakness	5.41	0.77	4.00
229.	Serious slipped disc	5.41	0.38	3.00
135.	Regular severe headache*	5.42	0.27	1.00
1.	Addiction to legal drugs (e.g. alcohol, nicotine, prescribed drugs)	5.44	0.77	3.00
32.	Malignant skin cancer-treatable with complete cure	5.44	1.09	4.00
57.	Severe concussion	5.44	0.90	3.00
216.	Chronic serious sleep disorder	5.44	0.45	3.00
6.	Occupant regularly violent when in building	5.47	0.58	3.00
14.	Constant serious anxiety	5.47	0.39	3.00
126.	Immediate mild heart attack	5.47	0.77	3.00
222.	Gradual permanent partial dependency	5.47	0.77	3.00
59.	Chronic severe wheezing/breathlessness (immediately induced)	5.50	0.84	4.00
190.	Chronic immediate severe personality change	5.50	0.97	3.00
270.	Serious ligament injury in neck*	5.58	0.27	1.00
19.	Chronic severe backache	5.59	0.51	3.00
30.	Serious chronic behaviour disorder (immediately following hazardous event)	5.59	0.44	3.00
103.	Regular severe fever	5.59	0.57	4.00
226.	Mild stroke	5.62	0.69	4.00
239.	Chronic/regular serious infection of lung	5.62	1.15	5.00

HARM		MEAN	VAR.	RANGE
117.	Simple fracture of legs	5.63	1.92	4.00
119.	Compound, complicated or comminuted fracture of ribs	5.66	0.56	3.00
129.	Chronic severe high blood pressure	5.66	0.36	3.00
115.	Simple fracture of back without damage to spinal cord	5.75	0.71	3.00
242.	Attempted suicide with full recovery	5.78	0.37	3.00
140.	Regular/chronic severe impairment of judgement	5.80	0.65	4.00
72.	Severe long-term depression	5.81	0.29	2.00
91.	Loss of consciousness for hours/days	5.81	0.74	3.00
125.	Multiple compound, complicated or comminuted fractures	5.81	0.35	3.00
9.	Long term amnesia*	5.83	0.15	1.00
77.	Chronic severe bronchitis*	5.83	0.15	1.00
210.	Severe permanently restricted child development (social/ emotional/cognitive)	5.87	0.65	2.00
187.	Permanent paralysis of leg	5.91	0.67	2.00
213.	Severe immunological sensitisation to chemical/biological agents	5.94	0.32	2.00
25.	Loss of foot	6.00	0.58	2.00
134.	Severe hypothermia	6.00	0.77	4.00
22.	Loss of dominant hand	6.03	0.61	2.00
2.	Use of class a illegal drugs	6.09	0.67	3.00
137.	Immediate permanent deafness in both ears	6.12	0.89	5.00
106.	Gangrene in upper limb	6.13	1.02	4.00
221.	Immediate permanent complete dependency	6.13	0.69	3.00
229.	Chronic severe psychosis*	6.17	0.15	1.00
245.	Severe retinal haemorrhage*	6.25	0.20	1.00
105.	Gangrene in hand or foot	6.28	0.60	3.00
55.	Minor immediate brain damage*	6.33	0.24	1.00
157.	Severe angina*	6.33	0.24	1.00
207.	Severe birth defect	6.34	0.56	2.00
65.	Severe chronic confusion/dementia	6.41	0.70	3.00
84.	Serious dislocated spine*	6.42	0.27	1.00
105.	Severe electric shock*	6.42	0.27	1.00
20.	Loss of arms	6.44	0.25	1.00
55.	80% burns/scalds	6.50	0.77	3.00
195.	Regular severe pneumonia	6.50	0.45	2.00
57.	Moderate gradual brain damage*	6.58	0.27	1.00
261.	Serious stroke*	6.67	0.24	1.00
33.	Malignant lung cancer-treatable with risk of recurrence or development of secondaries	6.72	0.34	2.00
34.	Malignant skin cancer - non-treatable	6.72	0.47	2.00
39.	Immediate permanent blindness in both eyes	6.72	0.21	1.00
64.	Severe burns/scalds to face*	6.75	0.20	1.00
125.	Regular grand mal seizure*	6.75	0.20	1.00
152.	Delayed severe heart attack*	6.75	0.20	1.00
240.	Fatal birth defect*	6.83	0.15	1.00
141.	Fracture of neck with damage to spinal cord*	6.83	0.15	1.00
21.	Loss of arms and legs	6.84	0.14	1.00
188.	Permanent paralysis below neck	6.91	0.09	1.00
16.	Asphyxiation resulting in major brain damage/death	7.00	0.00	0.00
35.	Leukaemia - non-treatable	7.00	0.00	0.00
44.	Fatal immediate brain damage	7.00	0.00	0.00
62.	Acute fatal bronchitis	7.00	0.00	0.00
86.	Fatal electric shock	7.00	0.00	0.00
92.	Permanent loss of consciousness	7.00	0.00	0.00
132.	Fatal hyperthermia	7.00	0.00	0.00
170.	Fatal internal abdominal damage	7.00	0.00	0.00
227.	Fatal stroke/heart attack	7.00	0.00	0.00
243.	Suicide	7.00	0.00	0.00

APPENDIX E.
COMPARISON OF EXPERT AND PUBLIC RATINGS OF SEVERITY OF HARM

E1. INTRODUCTION

The work described here is concerned with the comparison of ratings of severity of harm made by experts and the general population. Comparison of expert and lay perceptions of risk, for example in the fields of toxicology or nuclear waste, often highlights differences (see Kraus et al 1992, Flynn et al 1993). This, however, may not be the case when dealing with ratings of the severity component of risk, or when dealing with a scale composed of diverse harms. Any differences that do occur may, however, need to be considered in developing a scale of severity of harm.

The work was carried out in three phases, as described below: (i) a study of the rating of harms by a sample of the general public, using a subset of the harms used in the earlier study of experts (Appendix C), (ii) a study comparing results from a group of expert subjects, again using the subset of harms, with the results of the previous study which involved the full list of representative harms, and (iii) comparison of the expert and public ratings.

E2. PHASE ONE: PUBLIC RATINGS

E2.1 Introduction

This study, using the general public as subjects, involved a pilot study and a main study. The pilot study was carried out to test and refine the materials and procedures to be used in the main study. The main study was then carried out and the resulting data examined for sex and age differences and the possible influence that experience of a harm might have on its rating. Each harm was also assigned a severity score, based on its rating.

E2.2 Pilot Study

E2.2.1 Method

Nine male and 15 female subjects took part in the pilot exercise. The average age was 58 years, with range 23 years to 75 years. Subjects were canvassed by post from the local community in Loughborough, using a letter that briefly explained the purpose of the study and what participants would be asked to do. A reply slip and postage-paid envelope were provided.

A list of 64 harms was derived from those used in the main study of experts, as listed in Appendix D. Each fourth harm in this distribution was selected, and some harms were then replaced with adjacent harms to ensure broad outcome categories were represented. Each of these harms was printed on card (A6 size) and numbered on the reverse.

The seven point scale, used by subjects to rate the harms, was the same as that used in the main study of experts (Appendix C). The method was also as similar as possible, but adapted to be used in subjects' homes. The scale numbers and markers were used to label a set of A5 pockets. Each pocket was transparent on the front so that its contents could be seen during the sorting task. The eight pockets were attached to each other in a two by four array to form a 'sorting pack' which folded to A4 size. Instructions for the exercise were also attached to the pack; the wording was as follows.

> *You will be presented with a selection of around 65 harms. These items represent likely harms to building users.*
>
> *Your task is judge the severity of these items and sort the harms according to the following seven point scale:*
>
> *1 - Extremely Slight*
> *2 - Slight*
> *3 - Moderate/Slight*
> *4 - Moderate*
> *5 - Moderate/Severe*
> *6 - Severe*
> *7 - Extremely Severe*
>
> *Please study the full list of harms before beginning; this will familiarise you with the full range. Please place each card in the envelope appropriate to how severe you think it is. If you are uncertain of where a particular item fits the scale then it may be placed in the 'uncertain' pocket.*
>
> *Please consider each harm individually and in terms of what is presented on the card, and judge the harm's severity in terms of its overall impact on an individual's everyday life, e.g. how much quality of life would be affected, how long it would take to recover, etc. Please focus on the harm and not the possible cause, unless the cause is mentioned on the card. Some items are described as severe, or slight, or moderate, etc.; this does not necessarily mean that they should be placed on scale points with similar markers (for example slight heart attack and*

severe graze do not necessarily correspond to slight and severe on the scale).

After the initial sort you should try to place any cards you have put in the uncertain pocket on other points on the scale.

Subjects who replied to the postal survey were contacted to arrange a suitable time to call at their homes.

A pack of materials, consisting of the sorting pack, the 64 harm cards and a full list of items involved, was delivered to participants. Subjects were told that their task was to sort the harms according to their severity and place them in the appropriate scale pocket. The instructions were pointed out and arrangements made to call back for the completed pack.

When the materials were collected, subjects were asked (a) to estimate how long the sort took to complete, (b) if they had any general comments about the task, (c) if they had any specific comments about individual harms, and (d) if they had personal experience of any of the harms (that is, had they or someone close to them suffered any of the harms used).

E2.2.2 Results

The mean estimated time for completion of the whole task was about half an hour, with a range of 20 minutes to one hour. Three subjects commented that they found the task rather repetitive and tedious. Subjects made comments on the following aspects of the task.

- Four subjects commented that, while the clear front of the scale pockets allowed them to see which harms they had already placed in the pocket, it would be more useful to have the scale number and marker on the front so that they were visible throughout the sort.
- Several of the subjects reported that they had no idea what some of the harms were. This is reflected in the numbers of subjects placing harms in the uncertain category.
- Six of the subjects commented that they saw no point in including the descriptors ‘delayed' and ‘immediate' since this made no difference to their judgement of severity; they made that judgement based on the outcome ailment. Similarly the ‘acute' in ‘Acute fatal bronchitis' was felt to be irrelevant since ‘fatal' was the element judged to be crucial.

E2.3 Main study method

Seventy-seven subjects took part in the main card sort. These came from two sources. Four of the subjects replied to the original pilot study canvassing after the pilot study had been completed. The remaining subjects had replied as a result of the further postal canvassing of 2000 randomly selected addresses in Nottinghamshire, Leicestershire, Bedfordshire, Hertfordshire and Buckinghamshire.

The primary response from this mailshot was 247 replies (12.3%), of which 109 indicated that they would be willing to take part in the study. On contacting the subjects to arrange appointments a further 36 potential subjects now proved to be either unwilling to participate, unable to arrange an appointment or unobtainable. Of the subjects that took part, 34 were female and 43 male. The average age was 50 years (range 23-72).

The overall procedure for the main card sort remained the same as that of the pilot study but the following changes were made to the materials.

- The scale sorting pack was made more ‘user friendly' by moving the scale number and marker towards the top of the pocket.
- The most problematic individual items were those which could not be placed by several of the subjects in the pilot study. The following five items seemed the most problematic and were removed: *Temporary immediate slight personality change*, *Temporary moderate psychosis, Moderate chronic behaviour disorder (developing over time), Regular/chronic severe impairment of judgement* and *Severe permanently restricted child development (social/emotional/cognitive)*. The number of harms rated was thus reduced to 59.
- Finally, the adjectives that some subjects identified as redundant (for example immediate and delayed) were removed in an attempt to make the task clearer.

E2.4 Main study results

E2.4.1 Harm distribution

Each harm is shown in ascending order of severity in Table E1, with its mean rating. This distribution of harms is based on the combined responses from all 101 subjects taking part in the pilot and main sorts.

This distribution of severity ratings is skewed toward the severe end of the scale, as was the

case in the distribution in Appendix D, with over half of the harms rated above 4.5.

E2.4.2 Sex and age differences

It is possible that the subjects' age and sex may affect their perceptions of severity. Accordingly sex and age differences were tested for using a series of two-way ANOVA statistical tests on each harm. For the purposes of these tests, subjects' ages were grouped in three categories: under 35 years, 36 to 60 years, and over 60 years.

Two harms showed statistical differences for the sex of respondent (at the 0.05 level). Female respondents judged a *Strained Muscle/Tendon in Foot/Leg* more severe (mean 3.12) than male respondents (mean 2.54). Female respondents also judged *Addiction to Legal Drugs (e.g. Alcohol, Nicotine, Prescribed Drugs)* more severe (mean 5.56) than males (mean 5.07).

Five harms showed statistical differences for the age of respondent (again at the 0.05 level). Those in the lower age group judged *Occasional Slight Menstrual Disorder* as more severe (mean 2.80) than those in the middle age group (36 to 60) (mean 1.98). Respondents in the middle age group judged *Serious Punctures/Cuts to Arm/Wrist* as more severe (mean 5.64) than those in the upper age group (mean 4.73). Those in the middle age group also judged *Multiple Serious Punctures/Cuts* more severe (mean 5.74) than those in the lower age group (mean 5.09). This pattern is repeated with respect to *Regular Aggressive Behaviour When in Building*: those in the middle group judged this more severe (mean 5.78) than those in the lower age group (mean 4.88). Finally both those in the upper age group (mean 5.80) and those in the middle age group (mean 5.88) judged *Regular Severe Diarrhoea/Vomiting* as more severe than those in the lower age group (mean 5.21)

These results may be subject to Familywise Type I Error (see Keppel et al 1992) since such a large number of statistical tests were performed. With 59 harms and 3 comparisons per harm, a total of 177 tests were performed. One in 20 would yield a significant result in the absence of a real difference, thus 8.85 tests would be expected to pass the $p<0.05$ level of significance. Since 8 tests yielded a significant result there is no sound basis for accepting any of the significant differences. Another way of dealing with this is to adopt a stricter level of significance, for example 0.001. At this level no significant sex or age differences, or interactions between sex and age, emerged for any harm.

E2.4.3 Experience of harms

Just as subjects' sex and age may have affected their judgements of severity, so their experience of certain types of harm might also affect their perceptions of those types of harm. Subjects were asked if they had any personal experience of the harms involved in the study upon completion of the task. The numbers of subjects with experience of certain types of harm was seven with cardiovascular illness, two each with hearing difficulties and respiratory problems, and one each with cancer, stomach ulcer, vertigo, eye or sight problems and musculoskeletal problems.

Of these groups, only those with experience of cardiovascular illness form a large enough group to make any attempt to compare statistically with the rest of the sample. Ratings for three cardiovascular related harms (*Chronic Moderate High Blood Pressure, Mild Heart Attack* and *Short-Term Slight High Blood Pressure*) were compared across these two groups using independent sample t-tests. These tests showed no significant difference between the ratings of those subjects with experience of cardiovascular problems and those without. This is a very limited test and further studies would be necessary to show with confidence whether experience of similar harms, or more harms in total, affects ratings of severity.

E3. PHASE TWO: EXPERT RATINGS

E3.1 Introduction

The second study involved a group of expert subjects using the smaller sample of harms. This study allowed comparison of ratings with the results of previous studies based on the longer list of representative harms in Appendix D. This study also provides an expert comparison group for the study of the public.

E3.2 Method

The 25 subjects were drawn from two groups: (i) health and safety practitioners, and (ii) those with a medical background.

The same materials and procedure as used in Phase One were used in this study. The experts took the materials pack away, working on it in their own time and returning it the next day.

E3.3 Results

The mean estimated time for completion of the task was 20 minutes, similar to that for the general public.

Each harm is shown in Table E2 with its mean rating and range. Table E3 compares the means from Appendix D. In general the means are very similar; only three items are outside a 0.5 range of their original study values (*Strained Muscle/Tendon in Foot/Leg, Temporary Impaired Hearing in One Ear,* and *Serious Puncture/Cuts to Neck*) and these are within a range of 0.9. Those items at the 'extremely severe' end of the scale match completely. Thus the different procedure and smaller number of harms had a minimal effect on ratings, if any at all.

E4. COMPARISON OF PUBLIC AND EXPERTS

Each harm is shown in Table E3 with its associated mean rating from both the public and expert studies. The distributions are very similar with the experts' ratings of harms perhaps a little less skewed toward the severe end of the scale, suggesting that the general public may tend to overestimate severity relative to experts' judgements in some cases.

Differences between expert and public ratings of individual harms were tested using a series of independent samples t-tests. The mean ratings from each group for each item were compared, again using the 0.05 level to indicate a significant difference, although the possibility of Type I error should once again be noted. Of the 59 harms, 7 showed significant differences between expert and public ratings. The means for these harms are shown in Table E4: in each case the general public rated the harm as more severe than did the expert subjects.

It can be seen that, while the public ratings do differ from the expert ratings in the current study and are generally higher, they are more similar to those expert ratings derived from the original studies. Furthermore, given that 59 tests were performed, 3 significant differences would be expected to occur by chance in the absence of a real difference.

Any small difference in overall level of rating is, however, less important than a difference in the relative order of harms. To gauge this, the correlation between expert and public ratings was calculated and the degree of agreement was remarkable. The correlations between public and expert ratings was r = 0.990 (r = 0.984 for the original expert ratings). By comparison, the correlation between new and original expert ratings was r = 0.983).

E5. CONCLUSIONS

In general these studies have shown that there is little difference between expert and public ratings of severity of harm, and that these ratings are similar to those derived in earlier studies. This would seem to suggest that the scale derived in previous studies is relatively robust and forms a sound basis for the comparative evaluation of health and safety risks in buildings.

Although there are differences between the experts and public perceptions these would result in very few changes to the concluding tables in the original study Appendix D. Of the items rated by experts only three are outside a 0.5 range of the original study and only *Temporary Impaired Hearing in One Ear* would move from Class V to Class IV (as defined in Chapter 2). Of The significant differences between public and expert ratings, *Hairline Fracture of the Pelvis* would move from Class IV to Class III, and *Chronic Severe Stress* and *Regular Severe Diarrhoea/Vomiting* would move from Class III to Class II, based on the public ratings.

The differences are sufficiently small, even if statistically significant, that it would be safe to conclude that there is a general difference between ratings given by the general public and ratings given by experts. This contrasts with earlier studies of perceived risk, which generally do show a difference between experts and the public. This suggests that differences in perceived risk are due to differences in estimated likelihood or differences in the function used to combine severity and likelihood.

ACKNOWLEDGEMENT

The authors wish to acknowledge the assistance of Dan Rowles of BRE and Dr Luise Vassie of Loughborough University in the collection of data.

REFERENCES

Flynn, J, Slovic, P & Mertz, C K (1993) Decidedly different: Expert and public views of risks from a radioactive waste repository. Risk Analysis 13(6), 643-648.

Keppel, G, Saufley, W H & Tokunaga, H (1992) Introduction to Design and Analysis (2nd Edition). W.H. Freeman & Co: New York.

Kraus, N, Malmfors, T & Slovic, P (1992) Intuitive toxicology: Expert and lay judgements of chemical risks. Risk Analysis, 12 (2), 215-232.

Table E1. Ratings from main study of the public

HARM	MEAN	RANGE
Slight graze to hand	1.06	4
Occasional slight discomfort	1.22	3
Occasional slight headache	1.42	2
Occasional slight anxiety	2.17	4
Occasional slight irrational fear (e.g. of spiders enclosed spaces)	2.17	4
Occasional mild sleep deprivation	2.18	4
Temporary/occasional mild infection of ear	2.24	4
Occasional slight menstrual disorder	2.25	4
Occasional slight vertigo	2.46	5
Slight ligament injury in ankle	**2.50**	**4**
Moderate bruising to face	2.62	4
Strained muscle/tendon in foot/leg	2.67	4
Temporary slight wheezing/breathlessness	2.78	4
Regular slight dermatitis	2.80	4
Occasional irritability/intolerance towards other occupants	2.87	6
Short-term slight high blood pressure	2.89	4
Hairline fracture of nose	3.00	4
Temporary impaired hearing in one ear	3.14	4
Chronic/regular redness/rash other than on face or hands	3.46	5
Subclinical short-term depression	**3.58**	**6**
Temporary loss of hair because of disease	3.66	5
Regular moderate impairment of concentration	3.73	6
Regular moderate anxiety	3.89	5
Onset of mild asthmatic condition	3.89	5
Brief loss of consciousness	4.18	5
Short term amnesia	4.20	6
Occasional severe drowsiness	4.23	6
Benign tumours	4.35	6
Regular severe irrational fear (e.g. of spiders, enclosed spaces)	4.52	6
Regular severe loss of appetite	**4.57**	**6**
Regular severe lack of privacy	4.68	6
Mild chronic confusion/disorientation	4.76	5
Occasional moderate pneumonia	4.79	4
Acute moderate bronchitis	4.81	4
Stomach/duodenal ulcer	4.85	5
Hairline fracture of pelvis	4.87	4
Chronic moderate high blood pressure	5.00	5
Alienation from society (as an immediate result of occupying a certain home)	5.03	5
Serious puncture/cuts to arm/wrist	5.17	5
Severe short-term depression	**5.20**	**4**
Addiction to legal drugs (e.g. alcohol, nicotine, prescribed drugs)	5.40	5
Regular aggressive behaviour when in building	5.45	5
Severe tinnitus	5.50	5
Mild heart attack	5.50	4
Multiple serious puncture/cuts	5.51	5
Compound, complicated or comminuted fracture of ribs	5.54	5
Severe burns/scalds to hands	5.58	4
Regular severe fever	5.61	4
Chronic severe wheezing/breathlessness (immediately induced)	5.64	5
Serious puncture/cuts to neck	**5.64**	**4**
Regular severe diarrhoea/vomiting	5.73	3
Chronic severe stress	5.88	3
Severe hypothermia	6.19	3
Gangrene in upper limb	6.51	4
Severe chronic confusion/dementia	6.53	2
Malignant lung cancer: treatable with risk of recurrence or secondaries	6.81	2
Permanent paralysis below neck	6.91	4
Fatal internal abdominal damage	6.99	1
Fatal bronchitis	7.00	0

Every 10th harm is shown in bold in order to make the table more readable.

Table E2. Ratings from study of experts

HARM	MEAN	RANGE
Slight graze to hand	1.00	0
Occasional slight discomfort	1.16	1
Occasional slight headache	1.60	2
Occasional slight irrational fear (e.g. of spiders enclosed spaces)	1.96	2
Temporary/occasional mild infection of ear	2.12	2
Occasional slight anxiety	2.00	3
Occasional slight vertigo	2.20	2
Occasional slight menstrual disorder	2.35	3
Occasional mild sleep deprivation	2.04	3
Occasional irritability/intolerance towards other occupants	**2.44**	**3**
Slight ligament injury in ankle	2.20	3
Strained muscle/tendon in foot/leg	2.00	3
Moderate bruising to face	2.68	3
Hairline fracture of nose	2.68	3
Temporary slight wheezing/breathlessness	2.80	4
Regular slight dermatitis	3.04	3
Regular moderate impairment of concentration	3.40	5
Subclinical short-term depression	3.50	3
Short-term slight high blood pressure	2.52	4
Temporary impaired hearing in one ear	**2.80**	**4**
Regular moderate anxiety	3.76	4
Chronic/regular redness/rash other than on face or hands	3.56	4
Onset of mild asthmatic condition	3.60	4
Temporary loss of hair because of disease	3.80	4
Short term amnesia	3.84	4
Occasional severe drowsiness	3.59	4
Regular severe irrational fear (e.g. of spiders, enclosed spaces)	4.40	5
Benign tumours	4.56	6
Brief loss of consciousness	4.24	5
Hairline fracture of pelvis	**4.28**	**4**
Regular severe loss of appetite	4.44	4
Regular severe lack of privacy	4.24	3
Serious puncture/cuts to arm/wrist	4.80	2
Acute moderate bronchitis	4.60	4
Serious puncture/cuts to neck	5.48	4
Stomach/duodenal ulcer	4.80	4
Occasional moderate pneumonia	4.64	4
Mild chronic confusion/disorientation	4.08	5
Severe short-term depression	4.88	4
Multiple serious puncture/cuts	**5.20**	**5**
Alienation from society (as an immediate result of occupying a certain home)	4.75	5
Chronic moderate high blood pressure	5.08	5
Severe burns/scalds to hands	5.24	4
Regular aggressive behaviour when in building	5.00	3
Severe tinnitus	5.32	4
Chronic severe stress	5.52	4
Regular severe diarrhoea/vomiting	5.24	3
Addiction to legal drugs (e.g. alcohol, nicotine, prescribed drugs)	5.38	5
Mild heart attack	5.64	3
Chronic severe wheezing/breathlessness (immediately induced)	**5.54**	**4**
Regular severe fever	5.32	3
Compound, complicated or comminuted fracture of ribs	5.29	4
Severe hypothermia	5.68	5
Gangrene in upper limb	6.48	2
Severe chronic confusion/dementia	6.38	2
Malignant lung cancer: treatable with risk of recurrence or secondaries	6.68	2
Permanent paralysis below neck	6.72	2
Fatal bronchitis	7.00	0
Fatal internal abdominal damage	7.00	0

Table E3. Mean ratings by the public and experts

HARM	PUBLIC	EXPERTS	
		NEW	ORIGINAL
Slight graze to hand	1.06	1.00	1.06
Occasional slight discomfort	1.22	1.16	1.62
Occasional slight headache	1.42	1.60	1.92
Occasional slight anxiety	2.17	2.00	2.31
Occasional slight irrational fear (e.g. of spiders enclosed spaces)	2.17	1.96	2.09
Occasional mild sleep deprivation	2.18	2.04	2.50
Temporary/occasional mild infection of ear	2.24	2.12	2.25
Occasional slight menstrual disorder	2.25	2.35	2.42
Occasional slight vertigo	2.46	2.20	2.37
Slight ligament injury in ankle	**2.50**	**2.20**	**2.66**
Moderate bruising to face	2.62	2.68	2.88
Strained muscle/tendon in foot/leg	2.67	2.00	2.81
Temporary slight wheezing/breathlessness	2.78	2.80	3.06
Regular slight dermatitis	2.80	3.04	3.28
Occasional irritability/intolerance towards other occupants	2.87	2.44	2.63
Short-term slight high blood pressure	2.89	2.52	3.47
Hairline fracture of nose	3.00	2.68	3.00
Temporary impaired hearing in one ear	3.14	2.80	3.53
Chronic/regular redness/rash other than on face or hands	3.46	3.56	3.72
Subclinical short-term depression	**3.58**	**3.50**	**3.44**
Temporary loss of hair because of disease	3.66	3.80	3.91
Regular moderate impairment of concentration	3.73	3.40	3.38
Regular moderate anxiety	3.89	3.76	3.69
Onset of mild asthmatic condition	3.89	3.60	3.84
Brief loss of consciousness	4.18	4.24	4.25
Short term amnesia	4.20	3.84	4.00
Occasional severe drowsiness	4.23	3.59	4.03
Benign tumours	4.35	4.56	4.19
Regular severe irrational fear (e.g. of spiders, enclosed spaces)	4.52	4.40	4.13
Regular severe loss of appetite	**4.57**	**4.44**	**4.42**
Regular severe lack of privacy	4.68	4.24	4.47
Mild chronic confusion/disorientation	4.76	4.08	4.78
Occasional moderate pneumonia	4.79	4.64	4.77
Acute moderate bronchitis	4.81	4.60	4.53
Stomach/duodenal ulcer	4.85	4.80	4.72
Hairline fracture of pelvis	4.87	4.28	4.34
Chronic moderate high blood pressure	5.00	5.08	4.97
Alienation from society (as an immediate result of occupying a certain home)	5.03	4.75	4.91
Serious puncture/cuts to arm/wrist	5.17	4.80	4.50
Severe short-term depression	**5.20**	**4.88**	**4.81**
Addiction to legal drugs (e.g. alcohol, nicotine, prescribed drugs)	5.40	5.38	5.44
Regular aggressive behaviour when in building	5.45	5.00	5.19
Severe tinnitus	5.50	5.32	5.25
Mild heart attack	5.50	5.64	5.47
Multiple serious puncture/cuts	5.51	5.20	4.81
Compound, complicated or comminuted fracture of ribs	5.54	5.29	5.66
Severe burns/scalds to hands	5.58	5.24	5.16
Regular severe fever	5.61	5.32	5.59
Chronic severe wheezing/breathlessness (immediately induced)	5.64	5.54	5.50
Serious puncture/cuts to neck	**5.64**	**5.48**	**4.63**
Regular severe diarrhoea/vomiting	5.73	5.24	5.41
Chronic severe stress	5.88	5.52	5.34
Severe hypothermia	6.19	5.68	6.00
Gangrene in upper limb	6.51	6.48	6.13
Severe chronic confusion/dementia	6.53	6.38	6.41
Malignant lung cancer - treatable with risk of recurrence or secondaries	6.81	6.68	6.72
Permanent paralysis below neck	6.91	6.72	6.91
Fatal internal abdominal damage	6.99	7.00	7.00
Fatal bronchitis	7.00	7.00	7.00

Table E4. Comparison of public and expert mean ratings

HARM	PUBLIC	EXPERT	
		NEW	ORIGINAL
Strained muscle/tendon in foot/leg***	2.67	2.00	2.81
Occasional severe drowsiness*	4.23	3.84	4.03
Hairline fracture of pelvis**	4.87	4.28	4.34
Mild chronic confusion/disorientation**	4.76	4.08	4.78
Chronic severe stress*	5.88	5.52	5.34
Regular severe diarrhoea**	5.73	5.24	5.41
Severe hypothermia**	6.19	5.68	6.00

Significance of difference between public and experts in this study: *0.05; **0.01; ***0.001

APPENDIX F.
RATING OF HAZARDS

F1. INTRODUCTION

This study aimed to examine how experts and non-experts rate a list of hazards (rather than harms, as in the previous studies), to determine:

- whether the rank order of hazards given by the risk assessment methodology (Chapter 4) accords with the rank order given independently by experts and non-experts in health and safety fields;
- whether experts and non-experts differ in their ratings; and
- the effect of the amount of information provided about the hazards on the rankings given by the two subject groups.

The primary purpose of the work was to evaluate and refine the risk assessment methodology. The work also contributes to our understanding of why experts and non-experts differ in their assessments of risk.

F2. SUBJECTS

The non-expert subjects were BRE employees and had no specialist knowledge of health and safety or medical issues but were expected to be knowledgeable on some aspects of buildings. In addition, nine subjects in the main study were recruited and tested in their homes in the Nottingham area.

Expert subjects were drawn from 1996 entry Occupational Health and Safety Management MSc students and Health and Safety Executive (HSE) staff attending residential courses at Loughborough University, and medical staff at Nottingham City Hospital. Wherever possible, subjects were tested while in residence at Loughborough University or in their own homes.

F3. PILOT STUDY

F3.1 Method

In the pilot study, 16 non-expert subjects took part (five female and 11 male, age range of 21 to 51 years). In addition, 24 expert subjects took part. None of the subjects had been involved in previous studies carried out by the researchers.

The subjects from each group were divided among three conditions:

1: those shown a completed risk matrix (see example in Section 4.1) for each hazard and told to take into account the number of people affected, the severity of the outcome and the strength of the evidence linking the outcome to the hazard;

2: those shown an empty risk matrix with explanatory text in boxes and told to take into account the number of people affected, the severity of the outcome and the strength of the evidence;

3: those not shown a risk matrix but advised to take into account the number of people affected, the severity of the outcome and the strength of the evidence.

The pilot study used a reduced list of hazards from Chapter 4, selected at random.

Hazards were augmented with explanations and a full list of the 32 hazards used is shown in Table F1. Each of the hazards was printed on A7 size card (together with the appropriate hazard matrix for Condition 1). For Conditions 1 and 2 an A4 sized empty risk matrix was also shown.

Subjects were informed verbally of the task and written instructions were also given, adapted to each of the three conditions. The main points of the instructions were as follows as follows.

> *Participants in previous research have allowed us to produce information on the nature of hazards in buildings. These 'hazards' affect both the health and safety of building users.*
>
> *You will be presented with a selection of around 30 hazards. These items represent likely hazards to building users in general. This includes all types of hazards in homes and commercial buildings (though for the purposes of this study, please consider non-domestic hazards as those pertaining to office buildings).*
>
> *Your task is to judge how minor or major you think these items are, and sort the 'hazards' according to the following seven point scale:*
>
> *1 2 3 4 5 6 7*
>
> *Very Minor ... Very Major*
>
> *Please study the full list of hazards and their explanations before beginning; this will allow you to see the full range.*

[Condition 1: Please also study the example risk matrix; these matrices allow us to estimate 'hazards' comparatively by taking into account the number of people affected, the severity of the outcome and the strength of available evidence. You should consider the matrices shown on each card when deciding where to place the ''hazards'.]

[Condition 2: Please also study the example risk matrix; these matrices allow us to estimate 'hazards' comparatively by taking into account the number of people affected, the severity of the outcome and the strength of available evidence. You should also take these factors into account when deciding where to place the 'hazards'.]

[Condition 3: You should take into account the number of people you judge to be affected, the severity of the outcomes and the strength of available evidence when making you decision on the 'hazard'. You should also take these factors into account when deciding where to place the 'hazard'.]

Please place each card in the envelope appropriate to how minor or major you think the hazard is. If you are uncertain of where a particular item fits the scale then it may be place in the 'uncertain' pocket.

Please consider each hazard individually and in terms of what is presented on the card and after the initial sort you should try to place any cards from the uncertain pocket on the scale.

A set of eight A5 size document pockets was supplied to the subjects; these represented the scale numbers (1 to 7) and the "uncertain" category. Each pocket was transparent so that the subject could see the contents during the task. The eight pockets were arranged in a two by four array which formed a "sorting pack" which could be folded to A4 size. The complete list of hazards was also attached to the pack.

When all hazards had been rated, subjects were asked to rank the hazards within each of the seven categories. Upon completion of the task, the time taken to complete the initial sort of hazards was recorded and subjects were asked: (i) to explain the basic strategies they used to decide on where a hazard should be placed on the scale; and (ii) if they had had difficulty in understanding any terms, etc.

F3.2 Results

The non-experts' mean completion time was 17.8 minutes with a range of 5 to 30 minutes; this was slightly longer than the experts required (mean 15 minutes, range 10 to 25 minutes). In both groups the timing varied between conditions, with Condition 1 generally taking slightly longer than the other conditions, probably as a result of more detailed information being given. The duration of the task was acceptable for both groups.

All the subjects reported that they had taken both severity of outcome and numbers of people affected into consideration.

Non-experts encountered two main problems were with the method. Firstly it was found that many of the subjects did not know whether they fully understood what was expected of them, even with the verbal and written instructions. This meant that several of the subjects did not know whether they had completed the task correctly or not. A second problem was of the wording and structure of the instructions, which were not considered to be particularly "user friendly". Refinements (detailed below) were therefore carried out before proceeding to the main study. Several of the subjects in Condition 2 could not understand the relevance of the empty risk matrix.

Two experts in Condition 3 noted that it was not always possible to judge numbers of people affected by the given hazard accurately and felt that some indication should be given.

F4. MAIN STUDY

F4.1 Method

There were 40 non-expert subjects in the main study, none of whom had taken part in the pilot study. Nineteen of these were male and 21 were female, with an age range of 21 years to 75 years. Forty-two expert subjects took part. The results for the experts were pooled with those obtained in the pilot study, making a total sample size of 66.

As a result of the problems encountered during the pilot study, several changes were made to the main study method for non-expert subjects.

(a) It was noted in the results from the pilot study that several of the subjects from Condition 2 could not understand the relevance of the empty risk matrix. Consequently an explanation of the degrees of severity of outcome was added in order to aid matrix understanding and relevance. An A4 sized table was also added to the instructions in each condition, which gave examples of harms that were relevant to each class of severity.

(b) The instructions to the method were refined in order to become more “user friendly” by altering the wording from "please consider non-domestic hazards as those *pertaining to* office buildings" to "please consider non-domestic hazards as those *occurring in* office buildings". Paragraphs were also split in order to make the text easier to read.

(c) The final improvement to the method involved the researcher sitting with the non-expert subjects throughout the completion of the task. This allowed queries with the methodology to be answered there and then, thus preventing confusion.

The instructions for the task were also altered to make it clear what subjects should do with cards placed in the “uncertain” point on the scale. Apart from these refinements, the same procedure was followed as in the pilot study.

F4.2 Results and discussion

F4.2.1 Non-experts: effects of conditions

The final distribution of hazards is presented in Tables F1-F3, in ascending order of severity for each condition. Generally the subjects made little use of the two extreme ends of the scale. This finding would have been most expected in Condition 3, where no added information was given to the subjects. However in Condition 1 it might have been expected that more of the scale would have been used as a result of the added information given, which would have permitted the subjects to make a more informed judgement and therefore use more of the scale.

In all three conditions the ranges were relatively high (the majority above 5) which would suggest that there was little consistency between subjects' ratings. The ranges for Condition 1 were slightly lower than for the other two conditions. This would suggest that the added information given in Condition 1 did permit the subjects to make a more informed judgement and consequently there was more consistency. However, standard deviations for each hazard in Conditions 1, 2 and 3 ranged from 0.85-2.21, 0.10-2.09 and 1.22-2.28 respectively, indicating that the ratings for each condition depended heavily on factors additional to the amount of information provided, such as personal knowledge.

Differences in ratings between conditions were tested using a series of 32 one-way ANOVAs. The hazards that gave significant results ($p<0.05$) are shown in Table F4. As so few of the tests gave significant results (6 out of the 32 hazards) it is possible to conclude that the varying degrees of information that were given in the different conditions did not greatly affect the subjects' ratings for the majority of the hazards.

It is interesting to note here that when these hazards are placed in rank order of means, the ranking follows that obtained that derived in Chapter 4, except for fungal growth (domestic) and explosions (domestic), which change relative places. This is not a major reversal since the risk indices for these two hazards were very close in the original scheme (63 and 66 respectively).

It should be noted that these results may be subject to Familywise Type I Error (Keppel et al, 1992). As 32 ANOVA tests were performed it is likely that one in twenty of these tests would give a significant result in the absence of a real difference. Consequently it is possible that one or two of these differences may be due to chance. A MANOVA test was attempted across all conditions and hazards but there were too few degrees of freedom to complete it.

F4.2.2 Experts: effects of conditions

The final distribution of the 32 hazards is presented in Table F5-F7, in ascending order of mean severity rating.

Differences between the ratings from each subject group were tested using a series of one-way ANOVA statistical tests. Half of the hazards showed some degree of significant difference and the results are summarised in Table F8. It should be noted, however, that these results may be subject to Familywise Type I Error, as explained in Section F4.2.1.

The differences between the ratings in Condition 1 and the other two conditions are not unexpected since more information was given, in the form of matrices, in that condition. The distributions of hazards from the three conditions would also suggest that the information given in Condition 1 encouraged subjects to use the whole of the scale with a wider range of ratings. A MANOVA test carried out across all conditions and hazards shows that there is indeed a significant difference ($p < 0.001$) between the conditions. This reinforces the conclusion that Condition 1 tends to differ from the other two conditions.

F4.2.3 Differences between subject groups

Differences between the ratings were tested using a series of 32 two-way ANOVAs. The hazards that gave statistically significant results are shown

in Table F9. Interactions signify that the differences between conditions depended on subject groups or vice versa - these have been identified using Tukey's post hoc test.

Nearly half the hazards (13 of 32) differed significantly, either between the groups overall or in some conditions (eight hazards differed significantly between experts and non-experts, and five gave interactions), suggesting that these groups did make different judgements. It should again be noted that these results may be subject to Familywise Type I Error.

F4.2.4 Correlations

The ranks for each hazard in each condition are shown Table F10. The comparison also shows that overall grouping of the hazards is broadly similar, with a few exceptions as noted above. Most of the non-expert rankings in Condition 1 are closer to the experts ranking, suggesting that the non-experts may have modified their evaluations when presented with more information. This can also be seen in the correlation coefficients shown in Tables F11 and F12.

Table F11 shows the rank correlation coefficients for each subject group and condition compared with each other and with the risk assessment procedure. All the correlations are statistically significant except that between non-experts in Condition 3 and the risk assessment method rankings. The correlations between direct ranking and risk assessment rankings increase with the amount of information given and the correlations are higher for experts than for non-experts.

There is a linear relationship among the three conditions for both experts and non-experts, and between the two subject groups for each condition (Table F12).

In the non-expert group, the strongest relationship occurs between Conditions 2 and 3 where the subjects were aware of the type of information that should be taken into account in rating but no actual information was given. Condition 1, where the most information was provided, is only slightly more strongly related to Condition 2 than to Condition 3. This suggests that the amount of information subjects were given made the most difference in Condition 1, rather than the matrix approach itself.

In the expert group, Condition 1 has a stronger relationship with Condition 2 than Condition 3, while Condition 3 has a stronger relationship with Condition 2 than it does with Condition 1. This reflects the amount of information that subjects were given in each condition.

The strongest relationship between expert and non-expert ratings is in Condition 1 where most information was given to both groups. Condition 2 shows a weaker relationship and Condition 3 shows the weakest of all.

F4.2.5 Comparison with risk assessment

The relative rankings derived from the studies described above were compared with the rankings derived in Chapter 4 using the risk assessment procedure (see Table F10). All rankings derived from the expert subject group in Condition 1 are within 5 places of the original ranking, with the exception of "sulphur dioxide and smoke (domestic)" (11 places), "house dust mites (non-domestic)" (9 places), "slips, trips and falls from buildings (non-domestic)" (8 places) and "noise (domestic)" (7 places).

The correlations between direct rankings and those derived from the original risk assessment procedure are shown in Table F11. Perhaps not surprisingly, all the rankings derived from the expert subject group under the three conditions correlate at a significant level with the original risk assessment procedure ranking, with a coefficient increasing with the amount of information given. The same is true for non-expert rankings, although these are of a lesser magnitude and those derived from Condition 3 do not give a significant correlation.

The fact that experts in the fields of health and safety and medicine (and to some extent non-experts) are using information on severity of harm, numbers of people affected and strength of evidence to produce such a similar ranking is promising for the risk assessment procedure.

F5. CONCLUSION

Although the task was complex, both experts and non-experts understood it sufficiently well to make an attempt, even though they had misgivings about whether they really understood the task or whether they had adequate information. This offers support for the approach of ranking hazards on the basis of the number of people affected and the severity with which they are affected, moderated by the degree of certainty about the hazard-harm relationship.

Experts in Condition 1, i.e. those shown the information given in the risk matrix, showed a remarkably high degree of agreement with the results of the risk assessment calculations ($r=0.917$). Non-experts' judgements were also moderately well correlated with the calculations ($r=0.690$). While the judgements made in this study cannot be used, strictly speaking, to

validate the risk assessment methodology, they do increase confidence in the procedure itself and in the degree of acceptance that its results are likely to achieve.

Within and between subject groups, reducing the amount of information available had the effect of reducing the correlation coefficients, so the effect of information availability is general, rather than specific to the comparison of experts and non-experts.

The lower correlations found where non-expert rankings were involved may be the result of these subjects being less sure of some of the hazards than the experts and consequently rating them differently: non-expert subjects seemed likely to score a hazard towards the centre of the scale if they were unsure of it.

Once they are in possession of risk information, in the form of risk matrices, non-experts are as similar to experts in possession of the same information ($r=0.775$) as two groups of experts without that information, i.e. experts in Conditions 2 and 3 ($r=0.756$). In other words, once the reliability of the ratings is considered, experts and non-experts agree as well as could reasonably be expected. This again increases confidence in the procedure but also implies that the reasons experts and non-experts often assess risks differently stem from differences in knowledge, rather than differences in combining different pieces of information.

REFERENCE

Keppel, G, Saufley, W H & Tokunga, H (1992) Introduction to design and analysis (2nd ed.). WH Freeman & Co, New York.

Table F1. Mean hazard ratings (non-experts, Condition 1)

HAZARD	MEAN	S.D.	RANGE
Oxides of nitrogen (ND)	2.00	1.65	1-6
Falling objects (D)	2.46	1.61	1-6
Collision/entrapment involving lifts (ND)	2.46	1.85	1-7
Collision/entrapment involving doors (D)	2.54	1.85	1-7
Particulates (D)	2.58	1.98	1-7
House dust mites (ND)	2.62	1.39	1-5
Falling objects (ND)	2.62	1.98	1-7
Collision/entrapment involving doors (ND)	2.62	2.02	1-7
Collision/entrapment involving windows (ND)	2.62	1.76	1-7
Fungal growth (ND)	2.69	1.32	1-5
Landfill gas (D)	3.31	2.21	1-7
Sanitary accommodation (ND)	3.54	1.51	2-7
Explosions (D)	3.62	1.50	2-7
Slips, trips and falls on the level (ND)	3.77	1.42	1-6
Slips, trips and falls on stairs and ramps (ND)	3.77	1.36	2-6
Slips, trips and falls involving baths (D)	3.85	1.07	3-6
Sources of infection other than sanitary accommodation (ND)	3.85	1.68	1-6
Fungal growth (D)	3.92	1.12	2-6
Electrical hazards (ND)	4.00	1.83	2-7
Noise (D)	4.08	1.66	2-7
Sanitary accommodation (D)	4.08	1.04	2-6
Slips, trips and falls from buildings (ND)	4.08	1.44	2-7
Radon (ND)	4.23	1.74	2-7
Environmental tobacco smoke (ND)	4.31	1.18	2-6
Slips, trips and falls on stairs and ramps (D)	4.38	1.26	2-6
Electrical hazards (D)	4.38	1.33	3-7
Carbon Monoxide (ND)	4.54	1.56	2-7
Lead (D)	4.69	1.38	3-7
Burn/scalds hazards (D)	5.31	0.85	4-7
Sulphur dioxide and smoke (D)	5.31	1.55	3-7
Hygrothermal conditions (D)	5.54	1.13	4-7
Carbon Monoxide (D)	5.92	1.26	3-7

D = Domestic; ND = Non-domestic

Table F2. Mean hazard ratings (non-experts, Condition 2)

HAZARD	MEAN	S.D.	RANGE
Fungal growth (ND)	1.92	0.10	1-4
Fungal growth (D)	2.42	1.08	1-4
Collision/entrapment involving lifts (ND)	2.46	1.66	1-6
Noise (D)	2.58	1.62	1-6
House dust mites (ND)	2.62	1.50	1-5
Hygrothermal conditions (D)	2.82	1.66	1-6
Collision/entrapment involving windows (ND)	2.85	1.52	1-6
Sanitary accommodation (ND)	2.92	1.62	1-5
Collision/entrapment involving doors (ND)	3.00	1.35	1-6
Collision/entrapment involving doors (D)	3.08	1.55	1-6
Slips, trips and falls on the level (ND)	3.23	1.54	1-6
Sanitary accommodation (D)	3.33	1.57	1-6
Falling objects (D)	3.38	1.39	1-6
Particulates (D)	3.67	1.37	2-6
Lead (D)	3.92	1.38	2-6
Slips, trips and falls involving baths (D)	3.92	1.50	1-6
Sulphur dioxide and smoke (D)	3.92	2.02	1-7
Slips, trips and falls on stairs and ramps (ND)	3.92	1.80	1-7
Sources of infection other than sanitary accommodation (ND)	4.08	2.10	1-7
Falling objects (ND)	4.08	1.75	1-6
Oxides of nitrogen (ND)	4.09	1.81	1-6
Landfill gas (D)	4.23	2.09	2-7
Environmental tobacco smoke (ND)	4.23	1.79	1-7
Radon (ND)	4.33	2.06	1-7
Slips, trips and falls from buildings (ND)	4.46	1.45	2-7
Slips, trips and falls on stairs and ramps (D)	4.62	1.45	1-6
Electrical hazards (ND)	4.69	1.55	1-7
Burn/scalds hazards (D)	5.08	1.44	2-7
Electrical hazards (D)	5.23	1.24	2-7
Carbon Monoxide (ND)	5.38	1.98	1-7
Explosions (D)	5.54	2.03	2-7
Carbon Monoxide (D)	6.15	1.14	4-7

D = Domestic; ND = Non-domestic

Table F3. Mean hazard ratings (non-experts, Condition 3)

HAZARD	MEAN	S.D.	RANGE
House dust mites (ND)	2.71	1.33	1-5
Fungal growth (ND)	2.93	1.86	1-7
Collision/entrapment involving windows (ND)	3.00	1.52	1-6
Collision/entrapment involving doors (ND)	3.14	1.61	1-6
Fungal growth (D)	3.21	1.67	1-6
Collision/entrapment involving doors (D)	3.36	1.60	1-6
Falling objects (D)	3.36	2.10	1-7
Slips, trips and falls on the level (ND)	3.36	1.45	1-6
Hygrothermal conditions (D)	3.50	1.91	1-6
Sanitary accommodation (D)	3.50	1.91	1-7
Noise (D)	3.64	1.82	1-6
Sanitary accommodation (ND)	3.71	1.86	1-7
Collision/entrapment involving lifts (ND)	3.71	2.23	1-7
Lead (D)	3.93	2.16	1-7
Falling objects (ND)	3.93	2.06	1-7
Slips, trips and falls on stairs and ramps (ND)	4.14	1.75	1-7
Sulphur dioxide and smoke (D)	4.15	1.82	2-7
Environmental tobacco smoke (ND)	4.36	1.45	2-6
Slips, trips and falls involving baths (D)	4.43	1.28	2-6
Radon (ND)	4.43	1.83	2-7
Sources of infection other than sanitary accommodation (ND)	4.50	1.65	1-7
Oxides of nitrogen (ND)	4.50	2.28	1-7
Particulates (D)	4.54	1.66	2-7
Slips, trips and falls from buildings (ND)	4.57	1.65	2-7
Slips, trips and falls on stairs and ramps (D)	4.64	1.22	3-6
Carbon Monoxide (ND)	4.69	1.97	1-7
Landfill gas (D)	4.79	2.12	2-7
Electrical hazards (ND)	4.79	1.48	2-7
Burn/scalds hazards (D)	5.07	1.73	2-7
Carbon Monoxide (D)	5.57	1.55	3-7
Electrical hazards (D)	5.64	1.28	3-7
Explosions (D)	5.64	2.24	1-7

D = Domestic; ND = Non-domestic

Table F4. Hazards (means) giving significant differences between the three conditions

HAZARD	CONDITION 1	CONDITION 2	CONDITION 3	F	P<
Electrical hazards (domestic)	4.38^{a}	5.23	5.64^{a}	3.36	0.05
Explosions (domestic)	$3.62^{a,b}$	5.54^{a}	5.64^{b}	4.49	0.02
Fungal growth (domestic)	3.92^{a}	2.42^{a}	3.21	3.97	0.03
Hygrothermal conditions (domestic)	$5.54^{a,b}$	2.82^{a}	3.50^{b}	9.61	0.00
Oxides of nitrogen (non-domestic)	$2.00^{a,b}$	4.09^{a}	4.50^{b}	5.72	0.01
Particulates (domestic)	2.58^{a}	3.67	4.54^{a}	4.19	0.02

The means sharing a superscript on any row were significantly different (Tukey tests)

Table F5. Mean hazard ratings (experts, Condition 1)

HAZARD	MEAN	S.D.	RANGE
Collision/entrapment involving lifts (ND)	1.64	0.73	1-4
Oxides of nitrogen (ND)	1.68	0.65	1-3
Particulates (D)	2.09	0.61	1-3
House dust mites (ND)	2.09	0.81	1-4
Collision/entrapment involving windows (ND)	2.09	0.92	1-4
Fungal growth (ND)	2.23	0.61	1-3
Landfill gas (D)	2.50	1.06	1-5
Sanitary accommodation (ND)	2.64	1.33	1-6
Collision/entrapment involving doors (ND)	2.73	0.94	2-5
Falling objects (ND)	2.77	1.69	1-7
Falling objects (D)	3.00	1.48	1-6
Collision/entrapment involving doors (D)	3.45	1.10	2-6
Noise (D)	3.73	1.08	1-5
Lead (D)	3.95	1.21	2-6
Sanitary accommodation (D)	4.00	1.35	1-6
Carbon Monoxide (ND)	4.00	1.23	2-6
Sulphur dioxide and smoke (D)	4.09	1.69	2-7
Electrical hazards (ND)	4.23	0.81	3-6
Fungal growth (D)	4.27	1.35	2-7
Sources of infection other than sanitary accommodation (ND)	4.41	1.30	1-6
Slips, trips and falls on the level (ND)	4.45	1.14	3-7
Explosions (D)	4.55	1.06	3-7
Slips, trips and falls involving baths (D)	5.00	1.11	2-7
Electrical hazards (D)	5.09	0.87	4-7
Radon (ND)	5.09	1.06	3-7
Slips, trips and falls from buildings (ND)	5.14	1.28	3-7
Slips, trips and falls on stairs and ramps (ND)	5.18	1.05	4-7
Environmental tobacco smoke (ND)	5.55	1.30	2-7
Burn/scalds hazards (D)	5.82	1.05	4-7
Carbon Monoxide (D)	6.18	0.59	5-7
Hygrothermal conditions (D)	6.23	0.75	4-7
Slips, trips and falls on stairs and ramps (D)	6.27	1.03	4-7

D = Domestic; ND = Non-domestic

Table F6. Mean hazard ratings (experts, Condition 2)

HAZARD	MEAN	S.D.	RANGE
Oxides of nitrogen (ND)	1.64	0.73	1-4
Collision/entrapment involving lifts (ND)	1.68	0.65	1-3
Collision/entrapment involving windows (ND)	2.09	0.61	1-3
Fungal growth (ND)	2.09	0.81	1-4
Fungal growth (D)	2.09	0.92	1-4
Collision/entrapment involving doors (D)	2.23	0.61	1-3
Collision/entrapment involving doors (ND)	2.50	1.06	1-5
Sanitary accommodation (ND)	2.64	1.33	1-6
House dust mites (ND)	2.73	0.94	2-5
Falling objects (D)	2.77	1.69	1-7
Sanitary accommodation (D)	3.00	1.48	1-6
Particulates (D)	3.45	1.10	2-6
Landfill gas (D)	3.73	1.08	1-5
Falling objects (ND)	3.95	1.21	2-6
Noise (D)	4.00	1.35	1-6
Lead (D)	4.00	1.23	2-6
Carbon Monoxide (ND)	4.09	1.69	2-7
Sulphur dioxide and smoke (D)	4.23	0.81	3-6
Electrical hazards (ND)	4.27	1.35	2-7
Sources of infection other than sanitary accommodation (ND)	4.41	1.30	1-6
Slips, trips and falls on the level (ND)	4.45	1.14	3-7
Explosions (D)	4.55	1.06	3-7
Slips, trips and falls involving baths (D)	5.00	1.11	2-7
Electrical hazards (D)	5.09	0.87	4-7
Radon (ND)	5.09	1.06	3-7
Slips, trips and falls from buildings (ND)	5.14	1.28	3-7
Slips, trips and falls on stairs and ramps (ND)	5.18	1.05	4-7
Environmental tobacco smoke (ND)	5.55	1.30	2-7
Burn/scalds hazards (D)	5.82	1.05	4-7
Carbon Monoxide (D)	6.18	0.59	5-7
Hygrothermal conditions (D)	6.23	0.75	4-7
Slips, trips and falls on stairs and ramps (D)	6.27	1.03	4-7

D = Domestic; ND = Non-domestic

Table F7. Mean hazard ratings (experts, Condition 3)

HAZARD	MEAN	S.D.	RANGE
Noise (D)	2.14	1.13	1-5
Particulates (D)	2.20	1.20	1-4
Collision/entrapment involving doors (D)	2.27	1.12	1-6
Oxides of nitrogen (ND)	2.3	1.26	1-5
Fungal growth (ND)	2.55	1.10	1-4
Falling objects (D)	2.64	1.22	1-6
Lead (D)	2.73	1.49	1-6
Sanitary accommodation (D)	2.77	1.41	1-5
Sanitary accommodation (ND)	2.77	1.27	1-5
Fungal growth (D)	2.91	1.44	1-6
Sulphur dioxide and smoke (D)	3.00	1.88	1-7
Landfill gas (D)	3.00	1.65	1-6
House dust mites (ND)	3.14	1.49	1-6
Collision/entrapment involving windows (ND)	3.14	1.64	1-6
Falling objects (ND)	3.36	1.71	1-7
Collision/entrapment involving doors (ND)	3.41	1.53	1-6
Hygrothermal conditions (D)	3.68	1.46	1-6
Explosions (D)	3.91	1.69	1-7
Collision/entrapment involving lifts (ND)	3.91	1.66	1-7
Radon (ND)	4.10	2.07	1-7
Sources of infection other than sanitary accommodation (ND)	4.14	1.46	1-7
Environmental tobacco smoke (ND)	4.27	1.39	2-6
Electrical hazards (ND)	4.32	1.39	2-7
Carbon Monoxide (ND)	4.45	1.92	1-7
Slips, trips and falls involving baths (D)	4.50	1.30	2-6
Slips, trips and falls from buildings (ND)	4.64	1.40	1-7
Slips, trips and falls on the level (ND)	4.68	0.99	2-6
Slips, trips and falls on stairs and ramps (ND)	5.09	1.38	2-7
Slips, trips and falls on stairs and ramps (D)	5.27	1.32	2-7
Electrical hazards (D)	5.36	1.00	3-7
Burn/scalds hazards (D)	5.50	1.30	3-7
Carbon Monoxide (D)	5.73	1.35	3-7

D = Domestic; ND = Non-domestic

Table F8. Significant differences between the three conditions

HAZARD	CONDITION 1	CONDITION 2	CONDITION 3	F	P<
Hygrothermal conditions (D)	$6.23^{a,b}$	$4.86^{a,c}$	$3.68^{b,c}$	23.13	0.00
Slips, trips and falls on stairs and ramps (D)	$6.27^{a,b}$	5.36^{a}	5.27^{b}	3.86	0.03
Fungal growth (D)	$4.27^{a,b}$	3.00^{a}	2.91^{b}	6.28	0.00
Noise (D)	3.73^{a}	3.73^{b}	$2.14^{a,b}$	12.02	0.00
Explosions (D)	4.55	5.27^{a}	3.91^{a}	4.90	0.01
Lead (D)	3.95^{a}	4.05^{b}	$2.73^{a,b}$	5.77	0.00
Sanitary accommodation (D)	4.00^{a}	3.27	2.77^{a}	3.81	0.03
Collision/entrapment involving doors (D)	3.45^{a}	3.00^{b}	$2.27^{a,b}$	6.85	0.00
Particulates (D)	2.09^{a}	$3.52^{a,b}$	2.20^{b}	10.84	0.00
Landfill gas (D)	2.50^{a}	3.68^{a}	3.00	3.62	0.03
Radon (ND)	5.09^{a}	3.52^{a}	4.10	5.08	0.01
Environmental tobacco smoke (ND)	$5.55^{a,b}$	4.59^{a}	4.27^{b}	4.71	0.01
Electrical hazards (ND)	4.23^{a}	5.05^{a}	4.32	3.59	0.03
House dust mites (ND)	$2.09^{a,b}$	3.14^{a}	3.14^{b}	5.55	0.01
Collision/entrapment involving windows (ND)	2.09^{a}	2.77	3.14^{a}	3.74	0.03
Collision/entrapment involving lifts (ND)	$1.64^{a,b}$	$2.55^{a,c}$	$3.91^{b,c}$	18.07	0.00

The means sharing a superscript on any row were significantly different (Tukey tests). D = Domestic; ND = Non-domestic.

Table F9. Significant differences between groups and interaction effects

SUBJECTS:	EXPERT			NON-EXPERT		
HAZARD ↓ GROUP:	1	2	3	1	2	3
DIFFERENCES BETWEEN EXPERTS AND NON-EXPERTS						
Slips, trips & falls on stairs/ramps (domestic)	6.3	5.4	5.3	4.4	4.6	4.6
Burn/scalds hazards (domestic)	5.8	5.9	5.5	5.3	5.1	5.1
Slips, trips & falls involving baths (domestic)	5.0	5.1	4.5	3.9	3.9	4.4
Slips, trips & falls on the level (non-domestic)	4.5	4.5	4.7	3.8	3.2	3.4
Slips, trips & falls on stairs & ramps (non-domestic)	5.2	5.0	5.1	3.8	3.9	4.1
Slips, trips and falls from buildings (non-domestic)	5.1	5.1	4.5	4.1	4.5	4.6
Sulphur dioxide and smoke (domestic)	4.1	3.9	3.0	5.3	3.9	4.2
Landfill gas (domestic)	2.5	3.7	3.0	3.3	4.2	4.8
INTERACTIONS BETWEEN GROUPS AND CONDITIONS						
Hygrothermal conditions (domestic)	6.2^{abcd}	4.9^{aefg}	3.7^{beh}	5.5^{hij}	2.8^{cfi}	3.5^{dgj}
Noise (domestic)	3.7^{a}	3.7^{b}	2.1^{abcd}	4.1^{c}	2.6	3.6^{d}
Explosions (domestic)	4.5	5.3^{ab}	3.9^{acd}	3.6^{bef}	5.5^{ce}	5.6^{df}
Particulates (domestic)	2.1^{abc}	3.5^{ad}	2.2^{def}	2.6^{g}	3.7^{be}	4.5^{cfg}
Oxides of nitrogen (non-domestic)	1.7^{ab}	2.1^{cd}	2.3^{ef}	2.0^{gh}	4.1^{aceg}	4.5^{bdfh}

Means sharing a superscript on any row were significantly different ($p<0.05$) using Tukey's post hoc test.

Table F10. Rankings given by experts and non-experts

GROUP:	EXPERT				NON-EXPERT		
CONDITION:	ORIGINAL	1	2	3	1	2	3
Hygrothermal conditions (D)	1	2	11	16	2	27	23
Slips, trips and falls on stairs and ramps (ND)	2	1	4	4	7	7	8
Burns/scalds hazards (D)	3	4	2	2	3	5	4
Slips, trips and falls from buildings (ND)	4	12	13	6	18	22	25
Carbon monoxide (D)	5	3	1	1	1	1	3
Radon (ND)	5	8	20	13	10	9	13
Slips, trips and falls on the level (ND)	7	6	10	5	18	15	17
Electrical hazards (D)	8	8	3	3	7	4	1
Environmental tobacco smoke (ND)	8	5	12	11	9	10	15
Fungal growth (D)	10	14	27	23	15	31	28
Slips, trips and falls involving baths (D)	10	7	7	7	11	8	9
Sources of infection other than sanitary accommodation (ND)	10	13	14	12	16	13	11
Electrical hazards (ND)	13	15	9	10	14	6	5
Noise (D)	13	20	17	32	11	29	22
Slips, trips and falls on stairs and ramps (D)	13	10	6	8	16	15	13
Explosions (D)	16	11	5	15	20	2	1
Carbon monoxide (ND)	17	17	7	9	6	3	7
Lead (D)	17	19	15	26	5	15	18
House dust mites (ND)	19	28	24	19	24	28	32
Sanitary accommodation (D)	19	17	22	24	11	21	23
Collision/entrapment involving doors (D)	21	21	27	30	29	23	25
Falling objects (D)	22	22	23	27	30	20	25
Falling objects (ND)	22	23	17	18	24	13	18
Collision/entrapment involving doors (ND)	24	24	25	17	24	24	29
Fungal growth (ND)	25	27	29	28	23	32	31
Particulates (D)	25	28	20	31	28	19	12
Sanitary accommodation (ND)	27	25	25	24	21	25	20
Sulphur dioxide and smoke (D)	27	16	16	21	3	15	16
Landfill gas (D)	29	26	19	21	22	10	5
Collision/entrapment involving windows (ND)	30	28	30	19	24	26	30
Collision/entrapment involving lifts (ND)	31	32	31	14	30	30	20
Oxides of nitrogen (ND)	32	31	29	21	32	12	11

D = domestic ND = non-domestic

Table F11. Rank correlation coefficients

	Experts			Non-experts		
Condition:	1	2	3	1	2	3
Experts (2)	0.82**	1				
Experts (3)	0.72**	0.76**	1			
Non-experts (1)	0.78**	0.72**	0.52**	1		
Non-experts (2)	0.54**	0.75**	0.61**	0.45*	1	
Non-experts (3)	0.46**	0.72**	0.53**	0.41*	0.91**	1
Risk assessment	0.92**	0.73**	0.66**	0.69**	0.37*	0.30

* p<0.05 ** p<0.01

Table F12. Correlation coefficients

	Experts			Non-experts		
Condition:	1	2	3	1	2	3
Experts (2)	0.84**	1				
Experts (3)	0.69**	0.82**	1			
Non-experts (1)	0.84**	-	-	1		
Non-experts (2)	-	0.77**	-	0.50**	1	
Non-experts (3)	-	-	0.66**	0.46**	0.92**	1

** p<0.01

APPENDIX G.
NUMERICAL WEIGHTING OF SEVERITY OF HARM

G1. INTRODUCTION

The risk assessment methodology described in Chapter 4 classifies harms in Classes I to IV, with weightings assigned to each Harm Class (as shown in Table G1), for the purpose of calculating a risk index. It was found, in a sensitivity analysis of these weightings, that hazards causing large numbers of less serious harms tended to move in relation to hazards causing a small number of more serious harms, resulting in uncertainty in the use of ranking. Additionally, no attempt had been made to assign different weights to different lives according to factors such as the age of the person affected. This study therefore aimed to increase confidence in the weightings given to the four Harm Classes and to obtain relative weightings for different age groups.

G2. SUBJECTS

The non-expert subjects were members of the general public. The majority were BRE staff as they had no specialist knowledge of health and safety or medical issues but were familiar with building issues. The remainder were members of the public from the Nottingham area. The experts in the health, safety and medical fields were drawn from 1996 entry Occupational Health and Safety Management MSc students and Health and Safety Executive (HSE) staff attending residential courses at the Centre for Hazard and Risk Management, Loughborough University, and medical staff at Nottingham City Hospital. Subjects took part in this study after studies of rating hazards (Appendix F), and rating the most severe harms (Appendix I).

G3. PILOT STUDY

G3.1 Method

Eight non-expert and seven expert subjects took part in the pilot study. They were shown a table that illustrated the four different classes of harm (as in Table G1) and were then asked to assign values to these classes of harm, given that one million pounds would avoid ten cases of the most severe harms (Class I harms). Subjects were asked to decide how many harms in the other classes should be avoided by spending the same amount of money as was spent to avoid the ten Class I harms.

Once this part of the task had been completed, the subjects were asked to repeat the same procedure but for different age groups (under 18, 18-40, 40-60 and over 60). This part of the study was carried out by first selecting one age band at random for each subject and asking the subject to complete the same task given that the people involved were of that age group only. Once the subject had completed this part, s/he was asked to provide the same, or different values for the other three age bands.

The materials were left with the subjects and a time was arranged for collection. Subjects were informed of the basic instructions verbally, with more detailed instructions supplied in writing.

Upon completion of the whole task, subjects were asked: (i) to explain the basic strategies they had used to decide on values, (ii) if they had had difficulty in understanding the task or any of the terms involved, and (iii) if they had experience of any of the harms or outcome groupings used.

G3.2 Results

Most subjects said that they had applied some sort of multiplication or division of the initial sum and had tried to apply this throughout the task so that value for Class IV was twice (or 100 time etc,) as much as that for Class III and so on.

Three of the expert subjects commented that they had had experience of some of the harm outcome groupings involved in the example table but all said that this had not affected their decisions.

Several of the non-expert subjects had difficulties in understanding the task. These subjects did not know whether they understood what was expected from them and so consequently did not know whether they had completed the task correctly or not. One subject did not complete the second task of assigning weightings for the age groups. A second problem was with the wording of the written instructions, which were considered to be not particularly "user friendly".

G4. MAIN STUDY

G4.1 Method

Forty non-expert subjects took part in the main study. Nineteen of these were male and 21 were female with age range from 21 years to 75

years. Twenty expert subjects took part, in addition to the seven who took part in the pilot study.

The problems reported in the pilot study resulted in improvements in the clarity of the instructions. As several of the subjects did not fully understand what was expected from them in the pilot study it was decided that in the main study the researcher should sit with each non-expert subject throughout the completion of the task. This allowed queries with the methodology to be answered there and then, preventing confusion.

Apart from these refinements, the method used in the main study followed the same procedure as the pilot study.

G4.2 Results and discussion

Arithmetic mean scores for each age group and each class of harm are shown in Table G2. The age-related figures are normalised to a value of 1 for Class I in the under 18 age group. Arithmetic means can be misleading where the distribution is skewed with a few extremely high values. Geometric means and medians were therefore also calculated (see Tables G3 to G5).

A three-way ANOVA (Subject Group x Harm Class x Age Group) test was performed on log-transformed data. This yielded one significant interaction, that for Subject Group x Harm Class, suggesting that experts and non-experts rated some classes of harm differently. The interaction was due primarily to experts' weightings increasing with the age of the persons affected while non-experts' weightings decreased with age. These opposing approaches to the treatment of age makes it difficult to make any recommendations for the inclusion of age as a factor in the risk assessment methodology, especially given the relatively small magnitude of the age effects.

The medians for the overall ratings given by the expert subject group, that is those experienced in the fields of health and safety, may be the most suitable for inclusion in a risk index calculation. These are shown in Table G6, inverted so that the more severe harms receive larger weightings.

G5. SUPPLEMENTARY STUDY

G5.1 Introduction

There was concern that subjects asked to weight classes of harm did not base their assessments on a detailed evaluation of the harms, but rather adopted a nominal multiplication factor between classes. This study sought to clarify this by altering the boundaries between classes and the number of classes in used.

G5.2 Method

Only expert subjects took part in the supplementary study. Three tasks were defined.

Task 1. Participants were asked to assign values to the four classes of harm, as in the main study. This meant considering how many incidences of harms from each class are equally acceptable to 10 harms from Class I, in terms of funds spent to avert them.

Task 2. The same as Task 1 except there were six classes of harm, rather than four. The six classes were devised by splitting each of the original Classes II and III into two classes, as in Table G7.

Task 3. The same as Task 2 except that subjects were asked to consider how many incidences of harms from each class are equally acceptable to 1 (rather than 10) harms from Class I.

Subjects were divided randomly into three groups, each assigned tasks as follows.

Group 1. Task 1, followed by Task 2 (at least two days later).

Group 2. Task 2, followed by Task 1 (at least two days later).

Group 3. Task 3 only.

Tasks 1 and 2 were completed by 35 subjects and Task 3 by 28 subjects.

G5.3 Results and discussion

Aggregate weightings are shown in Tables G8 and G9. The means are shown for information but are highly influenced by extreme scores; the medians are therefore a better measure.

Table G10 restates the median values, starting from a base of Class I = 1000. The figures indicate that the relationship between Classes I and II is relatively stable, and that the order of magnitude changes for the other relationships only in Task 3. While the figures indicate a need for further research, they do not give sufficient justification for any particular revised set of weightings.

The values in the main study do not vary with age group, except in the case of Class III, for which

the weighting varies from 20 to 50 (median of medians 35, mean of medians 31). It is also for Class III that there is least consensus between studies, although the variation is not of such a magnitude that it is likely to have a major impact on risk assessment outcomes.

Giving greater weight to the subjects in the main study (who carried out a range of tasks and therefore had more chance to become familiar with the concepts, on balance a figure of 30 is suggested for Class III.

G7. CONCLUSION

Comparison of expert and non-expert severity weightings showed that there were differences between the two groups. The weightings also differed between the three measures of central tendency used (arithmetic mean, geometric mean and median). For example the ratio between Class I and Class IV could be 50 or 10000, depending on which results are regarded as most valid and reliable. None of the figures comes close to the 100,000 used in the original risk assessment methodology.

Values (based on the median value) of 1, 10, 30 and 1000 for Classes IV (least severe) to I (most severe) respectively are suggested.

Subjects on the whole did not greatly alter weightings when applying them to different age groups, suggesting that they did not value lives differently for different age groups.

Table G1. Harm Classes with weightings and examples of harms

Harm Class	Weighting	Example harms
Class I (extremely severe)	100,000	Death, permanent paralysis below neck, permanent loss of consciousness, regular severe pneumonia, 80% burns & scalds,
Class II (severe)	2,500	Severe chronic confusion/dementia, chronic severe backache, loss of foot, mild stroke, regular severe fever
Class III (moderate/severe)	50	Chronic immediate severe stress, regular severe migraine, occasional moderate pneumonia, severe burns/scalds to hands
Class IV (moderate)	1	Occasional severe discomfort, regular serious cough/cold, brief loss of consciousness, benign tumours, 10% burns/scalds

Table G2. Arithmetic means of severity weightings

		Harm Class			
Age of persons affected	Subject group	IV	III	II	I
Not specified	Experts	2625	168	9.0	1
	Non-Experts	2091	147	16.7	1
Under 18	Experts	2626	120	7.9	1
	Non-Experts	2406	153	23.1	1
18 - 40	Experts	2676	122	12.1	1.4
	Non-Experts	2040	124	20.6	1.1
40 - 60	Experts	2756	139	10.6	1.8
	Non-Experts	1651	134	20.0	1.2
Over 60	Experts	2941	367	25.7	4.8
	Non-Experts	1930	208	22.0	1.7

Table G3. Geometric means of severity weightings

Age of persons affected	Subject group	Harm Class			
		IV	III	II	I
Not specified	Experts	529	48.5	6.3	1.00
	Non-Experts	89	26.3	6.1	1.00
Under 18	Experts	462	41.1	5.7	1.00
	Non-Experts	73	26.9	6.8	1.00
18 - 40	Experts	507	46.6	6.9	1.11
	Non-Experts	65	21.7	5.6	0.95
40 – 60	Experts	526	50.4	6.8	1.15
	Non-Experts	65	19.8	5.9	0.88
Over 60	Experts	612	68.1	9.2	1.25
	Non-Experts	58	19.3	5.2	0.83

Table G4. Geometric means of severity weightings, normalised to Class I

Age of persons affected	Subject group	Harm Class			
		IV	III	II	I
Not specified	Experts	529	48.5	6.3	1
	Non-Experts	89	26.3	6.1	1
Under 18	Experts	462	41.1	5.7	1
	Non-Experts	73	26.9	6.8	1
18 - 40	Experts	457	42.0	6.2	1
	Non-Experts	68	22.8	5.9	1
40 - 60	Experts	457	43.8	5.9	1
	Non-Experts	74	22.5	3.3	1
Over 60	Experts	490	54.5	7.4	1
	Non-Experts	70	23.3	6.3	1

Table G5. Medians of severity weightings, normalised to Class I

Age of persons affected	Subject group	Harm Class			
		IV	III	II	I
Not specified	Experts	1000	28	10	1
	Non-Experts	50	11	5	1
Under 18	Experts	1000	50	10	1
	Non-Experts	60	30	7.5	1
18 - 40	Experts	1000	25	10	1
	Non-Experts	25	12.5	5	1
40 - 60	Experts	1000	32	10	1
	Non-Experts	30	10	5	1
Over 60	Experts	1000	20	10	1
	Non-Experts	20	12	4	1

Table G6. Medians of weightings given by experts, normalised to Class IV

Harm Class	I	II	III	IV
Median	1000	100	36	1

Table G7. New Harm Classes

CLASS	EXAMPLES				
IV	Occasional severe discomfort	Moderate ligament injury in hip	Regular serious cough/cold	Benign tumours	Brief loss of consciousness
IIIb	Severe short-term depression	Loss of finger	Occasional moderate pneumonia	Stomach/ duodenal ulcer	Chronic moderate high blood pressure
IIIa	Severe burns/scalds to hands	Temporary paralysis of arms	Immediate mild heart attack	Serious slipped disc	Regular severe migraine
IIb	Severe long-term depression	Chronic severe backache	Mild stroke	Regular severe fever	Loss of consciousness for hours/days
IIa	Severe chronic confusion/ dementia	Loss or arms	Severe hypothermia	Gangrene in hand or foot	Immediate permanent deafness in both ears
I	Suicide	Loss of arms and legs	Fatal stroke/ heart attack	Malignant lung cancer	Permanent loss of consciousness

Table G8: Mean weightings

CLASS	TASK 1	TASK 2	TASK 3
IV	1.00	1.00	1.00
IIIb		12.1	12.5
IIIa	15.2	71.0	64.8
IIb		299	283
IIa	219	1300	3040
I	2600	9660	63100

Table G9: Median weightings

CLASS	TASK 1	TASK 2		TASK 3	
IV	1	1	1	1	1
IIIb		10		10	
IIIa	10	20	15	20	15
IIb		100		100	
IIa	100	200	150	1000	550
I	1000	1000	1000	10000	10000

Table G10. Median weightings with Class I = 1000

CLASS	MAIN STUDY	TASK 1	TASK 2	TASK 3	TASKS 2 &3
I	1000	1000	1000	1000	1000
II	100	100	150	55	102.5
III	36	10	15	1.5	8.25
IV	1	1	1	0.1	0.55

APPENDIX H.
NUMERICAL WEIGHTING OF STRENGTH OF EVIDENCE

H1. INTRODUCTION

The risk assessment methodology described in Chapter 4 assigns numerical weightings to the strength of the evidence for a link between a hazard and the number of people affected by a given Harm Class (see Chapter 3). The strength of evidence was classified as follows:

High (***)	a link between hazard and outcome is well established by firm evidence;
Medium (**)	probably an effect but less complete evidence is available;
Low (*)	possibly a link between hazard and outcome.

These strengths of evidence were assigned values of 5, 3 and 1 respectively in the calculation of a risk index. The aim of the study was to increase confidence in these values by using costs of avoidance as an index, based on judgements made by experts and non-experts.

The subjects were the same as those used in the study assigning weighting to severity classes (see Appendix G). Expert subjects were given the two tasks separated, in most cases by a period of a few days. Non-expert subjects undertook both tasks in the same session.

H2. PILOT STUDY

H2.1 Method

Eight non-experts and eight experts took part in the pilot study. The method was essentially the same procedure as the severity weighting study except that the three levels of strength of evidence were compared and age groups were not separated.

The eight subjects in this pilot study were given basic instructions verbally and more detailed instructions in written form. The written instructions provided an explanation of different categories of strength of evidence along with some examples to illustrate the different classes of harm. The subjects were then asked to assign a value to each level of evidence for each class of harm, given that one million pounds (£1,000,000) would certainly avoid a very severe harm (Class I). Thus, the larger the number given by the subjects, the more severe the combination of evidence and Harm Class was considered to be. Subjects were asked to consider the source and nature of the evidence and their perceptions of its "reliability" in their response.

The materials were left with the subjects and a time arranged for collection. At the collection time the subjects were asked (i) to explain the basic strategies they used to decide their values; and (ii) if they had any difficulty in understanding any terms, etc.

H2.2 Results and discussion

The subjects reported that they had taken both the class of harm and the strength of evidence into consideration in assigning values. Several of the subjects noted that they used some sort of multiplication or division in order to obtain their values (for example, the value for high strength of evidence was twice that for medium strength of evidence).

Two problems were encountered by non-experts with the pilot study method. Firstly it was found that the instructions were not considered to be particularly "user friendly". Secondly, several subjects did not know if they understood what was expected of them. This meant that they did not know whether they had completed the task correctly or not. Most expert subjects also noted that this exercise was conceptually difficult for them to grasp and required quite a lot of thought.

H3. MAIN STUDY

H3.1 Method

In the main study 40 non-expert subjects took part, 19 male and 21 female, with an age range from 21 years to 75 years. Twenty expert subjects took part.

As a result of the difficulties reported in the pilot study, two changes were made to the method for the main study. Firstly the instructions were modified in order for them to become more "user friendly". Secondly the researcher sat with each non-expert subject throughout the completion of the task. This allowed queries with the method to be answered there and then, thus preventing confusion. Otherwise, the method in the main study followed the same procedure as the pilot study.

H3.2 Results and discussion

Tables H1 to H3 show the arithmetic mean, geometric mean and median weightings respectively for each level of strength of evidence and each Harm Class for both expert and non-expert subjects.

The scores were normalised to a value of 1 for low strength of evidence in each Harm Class so comparisons cannot be made between Classes.

H4. CONCLUSION

Weightings given to strength of evidence were more consistent than those for Harm Class. However, as with Harm Class, the medians from the expert subject group may be the most appropriate to adopt for the high, medium and low values of strength of evidence since the medians are very similar across the four Harm Classes. These medians would suggest that values of 1, 2 and 4 be adopted for low, medium and high strengths of evidence compared to 1, 3 and 5 used in the original risk index methodology.

Table H1. Arithmetic mean weightings for each strength of evidence and Harm Class

SUBJECTS & STRENGTH OF EVIDENCE →	EXPERTS			NON-EXPERTS		
HARM CLASS ↓	HIGH	MEDIUM	LOW	HIGH	MEDIUM	LOW
IV	5.4	3.1	1	2511.2	7.0	1
III	6.6	3.6	1	9.0	3.4	1
II	54.9	27.5	1	9.2	3.2	1
I	5013.7	2507.3	1	25014.3	5004.1	1

Table H2. Geometric mean weightings for each strength of evidence and Harm Class

SUBJECTS & STRENGTH OF EVIDENCE →	EXPERTS			NON-EXPERTS		
HARM CLASS ↓	HIGH	MEDIUM	LOW	HIGH	MEDIUM	LOW
IV	4.2	2.5	1	5.1	2.5	1
III	5.1	2.9	1	3.8	2.2	1
II	6.0	3.2	1	3.6	2.1	1
I	7.2	4.6	1	4.8	2.8	1

Table H3. Median weightings given by experts

SUBJECTS & STRENGTH OF EVIDENCE →	EXPERTS			NON-EXPERTS		
HARM CLASS ↓	HIGH	MEDIUM	LOW	HIGH	MEDIUM	LOW
IV	4.5	2.3	1	3.2	1.8	1
III	4.5	2.0	1	3.3	1.9	1
II	4.0	2.0	1	2.6	1.5	1
I	4.0	2.2	1	2.0	1.6	1

APPENDIX I.
ASSESSMENT OF THE MOST SEVERE HARMS

I1. INTRODUCTION

In Chapter 2, subjects rated the severity of a list of harms directly against a seven point scale, from 1 (extremely slight) to 7 (extremely severe). The study resulted in 28 harms being placed within the 6-7 band. The aim of the present study was to rate only the harms in this most severe band in order to give a more precise comparison of harms within this band and thus provide greater confidence in the degrees of harm associated with specific health outcomes. In particular, the high degree of agreement between experts and non-experts found in Appendix E might be attributed to the major differences among the harms rated. If this is the case, then a lesser degree of agreement should be found when a narrower range of harms is assessed. The subjects were the same as those who had taken part in the study reported in Appendix F.

I2. PILOT STUDY

I2.1 Method

In the pilot study 12 non-expert subjects took part. Five were female and seven were male with an age range of 21 years to 51 years. Five expert subjects took part.

The subjects in the pilot study were shown a list of 38 harms. The majority of these harms appeared in the 6-7 band in Appendix D. The remainder were obtained from the full list in Appendix B and were selected for inclusion on the basis that they must logically be at least as severe as the harms rated 6.0 in the earlier study (e.g. fatal stroke/heart attack). The full list of 38 harms provided coverage of all of the broad outcome categories described in Table 3 of Chapter 3 (cardiovascular, respiratory etc.). The full list of harms used in the pilot study is shown in Table I1.

The subjects were asked to rate each of the severe harms against a seven point scale in which 1 equalled “least severe” and 7 equalled “most severe”. The method was otherwise as in Appendix C. The subjects were given basic instructions verbally, with more detail provided in the written instructions. The materials were left with the subjects and a time was arranged for collection.

At the collection time the subjects were asked (i) to explain the basic strategies they used to decide where each harm should be placed; and (ii) if they had any difficulty in understanding any of the terms, etc. The time taken to complete the sort was recorded.

I2.2 Results

The non-experts’ mean completion time for the whole task was 13.13 minutes with a range of 5 to 25 minutes. For experts the mean completion time was 11 minutes (range 8 to 20). The subjects reported that they had judged each harm’s severity in terms of its overall impact on the individual. They had also taken into account the effect of the harm on day-to-day living.

Non-experts encountered two problems with the method in the pilot study. The first concerned the understanding of the terminology of certain harms. Several subjects had difficulty in placing certain harms on the scale as they did not have enough knowledge about the severity of these harms in order to make a more informed judgement. The second problem was that some of the subjects did not know whether they fully understood what was expected of them, even with verbal and written instructions.

I3. MAIN STUDY

I3.1 Method

There were 40 non-expert and 18 expert subjects in the main study.

Although there was a problem in the pilot study in understanding the terminology of certain harms, it was decided that these harms would not be omitted from the main study. This was because the majority of the subjects *did* understand their meaning. It was decided, however, that the researcher should sit with the non-expert subjects throughout the completion of the task to obviate the problem experienced in the pilot study of not being sure about the instructions. This permitted queries with the methodology to be answered there and then, thus preventing confusion. Apart from these refinements, the same procedure was followed as in the pilot study.

I3.2 Results and discussion

The distributions of harm ratings are shown in Tables I1 and I2, in ascending order of severity. As in the pilot study the most severe end of the scale is characterised mainly by fatal harms.

In general, subjects tended not to use the lower end of the scale, perhaps not surprisingly since, by definition, these are the most severe of harms. For the experts, none of the mean rating scores is below 2.78 and the majority are above 4.50 (25 out of the 38 harms). For the non-experts, none of the mean rating scores is below 3.13 and, again, the majority are above 4.50 (25 out of the 38 harms).

The ranges of the scores for each harm are relatively large, indicating that there was either little agreement in general between the subjects or that there were a few extreme ratings. When the complete range of scores is examined, it is seen that most harms receive an even spread of scores along the scale, but for a few the range has been affected by extreme rating scores.

The harm *Fatal immediate brain damage* provides an example of the range score being affected by an extreme score. The mean score is 6.65 and most subjects rated this harm as 7. However scores of 1, 2, 5 and 6 were also recorded. It can be argued that the reason for this harm (and other harms with similar results) being rated low by some subjects is because the harm is perceived as having very little effect on an individual's subsequent quality of life as immediately fatal brain damage could be taken to imply that the individual will have no subsequent life.

In the pilot study it was noted that several subjects had difficulty in understanding certain harms. These subjects felt that they did not know enough about the severity of the harms to make a valued judgement. This was repeated in the main study. One type of harm which appeared to cause problems was psychological disorders, such as *Chronic delayed severe personality disorder* and *Serious chronic behaviour disorder (developing over time).* This may have led these harms to be given mean rating scores towards the middle of the scale (3.97 and 3.98 respectively). It could be that the subjects were not particularly sure where to place these harms, so in order to avoid making a judgement they placed the harms towards the middle of the scale.

If further research of this type were to be undertaken it may be important to consider whether explanations of the meanings of the harms should be included. This would aid the subjects in their understanding of the harms and consequently would enable them to place certain harms more easily on the scale. However it must also be considered that this may also *lead* the subjects in their answers since describing the harms may give indications (accurate or otherwise) about severity. Alternatively, the subjects could be asked to say how sure they were about the placing of each harm or to report how well they thought they understood the meaning of the harm.

There was a strong linear relationship between expert and non-expert ratings of severity of harm ($r=0.94$).

Differences between expert and non-expert ratings were tested using a series of independent samples t-tests. The mean ratings from each group were compared. Of the 38 harms, 4 showed significant differences ($p < 0.05$) between expert and non-expert ratings, although the possibility of Type I error should again be noted. The means for these harms are shown in Table I3. The differences are not all in one direction.

I4. CONCLUSION

Subjects assessed the most severe harms on the basis of the overall impact on the individual and the effect on day-to-day living. Fatal harms were rated as the most severe and the lower end of the severity scale tended to be avoided. Some subjects found difficulty in scaling harms through lack of knowledge of terminology and the severity of their effects.

The results of rating the most severe harms are very similar for the two subject groups and, of the 38 harms used, only 4 showed significant differences between expert and non-expert ratings. The differences are not all in one direction. The high degree of agreement between subject groups supports the use of the severity scale as derived in Appendix C.

Table I1. Harms in ascending mean order – non-experts

HARMS	MEAN	S.D.	RANGE
Chronic/regular serious infection of the skin	3.13	1.40	1-6
Chronic delayed severe stress	3.16	1.55	1-6
Use of Class A illegal drugs	3.31	1.85	1-7
Chronic/severe wheezing/breathlessness (gradually induced)	3.33	1.58	1-7
Loss of foot	3.50	1.40	1-7
Gangrene in hand or foot	3.51	1.48	1-7
Severe hypothermia	3.58	1.76	1-7
Regular severe pneumonia	3.85	1.53	1-6
Loss of dominant hand	3.95	1.08	2-6
Chronic delayed severe personality disorder	**3.97**	**1.63**	**1-7**
Serious chronic behaviour disorder (developing over time)	3.98	1.31	1-7
Immediate permanent deafness in both ears	4.18	1.22	2-7
Gangrene in upper limb	4.28	1.39	2-7
Compound, complicated or comminuted fracture of the skull	4.55	1.75	1-7
Severe chronic confusion/dementia	4.75	1.66	1-7
Malignant lung cancer – treatable with the risk of recurrence or secondaries	4.90	1.52	1-7
Severe birth defect	4.95	1.45	1-7
Loss of arms	4.98	0.86	3-6
Permanent paralysis of legs	5.00	1.18	3-7
Gradual permanent complete dependency	**5.05**	**1.14**	**3-7**
Immediate permanent blindness in both eyes	5.28	1.06	3-7
80% burns/scalds	5.48	1.30	1-7
Immediate severe heart attack	5.55	1.54	1-7
Immediate permanent complete dependency	5.68	1.14	3-7
Serious gradual brain damage	5.80	0.85	4-7
Malignant skin cancer – non-treatable	6.10	1.17	2-7
Loss of arms and legs	6.13	0.85	4-7
Permanent paralysis below neck	6.13	0.82	4-7
Suicide	6.26	1.73	1-7
Leukemia – non-treatable	**6.38**	**0.81**	**4-7**
Fatal internal abdominal damage	6.48	1.32	1-7
Permanent loss of consciousness	6.48	1.13	1-7
Fatal hyperthermia	6.53	1.32	1-7
Fatal electric shock	6.55	1.32	1-7
Acute fatal bronchitis	6.58	1.26	1-7
Fatal stroke/heart attack	6.63	1.25	1-7
Fatal immediate brain damage	6.65	1.25	1-7
Asphyxiation resulting in major brain damage/death	6.70	0.65	4-7

To improve readability, each 10th harm is printed in bold

Table I2. Harms in ascending mean order – experts

HARMS	MEAN	S.D.	RANGE
Chronic/regular serious infection of the skin	2.78	1.28	1-6
Use of Class A illegal drugs	3.30	1.15	1-5
Chronic/severe wheezing/breathlessness (gradually induced)	3.39	1.73	1-7
Loss of foot	3.87	0.92	2-6
Gangrene in hand or foot	4.04	0.93	1-5
Severe hypothermia	4.04	1.55	1-7
Severe chronic confusion/dementia	4.09	1.00	2-6
Chronic delayed severe personality disorder	4.13	1.10	2-6
Compound, complicated or comminuted fracture of the skull	4.13	1.14	2-6
Immediate permanent deafness in both ears	**4.17**	**0.98**	**2-6**
Loss of dominant hand	4.22	0.85	2-5
Chronic delayed severe stress	4.30	1.40	3-7
Serious chronic behaviour disorder (developing over time)	4.30	0.76	3-6
Gradual permanent complete dependency	4.52	1.24	1-6
Gangrene in upper limb	4.65	1.19	1-6
Regular severe pneumonia	4.65	1.19	2-7
Malignant lung cancer – treatable with the risk of recurrence of secondaries	4.70	1.18	2-6
80% burns/scalds	5.00	0.80	3-6
Permanent paralysis of legs	5.00	0.90	3-6
Severe birth defect	**5.05**	**1.13**	**3-7**
Loss of arms	5.09	1.08	2-6
Immediate permanent blindness in both eyes	5.13	0.81	4-6
Loss of arms and legs	5.13	1.01	2-6
Immediate permanent complete dependency	5.26	1.48	1-7
Permanent paralysis below neck	5.35	0.78	3-6
Malignant skin cancer – non-treatable	5.70	1.06	3-7
Immediate severe heart attack	5.83	0.78	4-7
Serious gradual brain damage	6.17	0.83	4-7
Suicide	6.22	1.13	2-7
Acute fatal bronchitis	**6.35**	**1.15**	**2-7**
Leukemia – non-treatable	6.35	0.88	4-7
Permanent loss of consciousness	6.52	0.95	3-7
Fatal electric shock	6.57	0.90	3-7
Fatal internal abdominal damage	6.61	0.78	4-7
Asphyxiation resulting in major brain damage/death	6.65	0.49	6-7
Fatal hyperthermia	6.83	0.49	5-7
Fatal immediate brain damage	6.83	0.49	5-7
Fatal stroke/heart attack	6.83	0.39	6-7

To improve readability, each 10th harm is printed in bold

Table I3. Comparison of expert and non-expert mean ratings

HARM	EXPERT	NON-EXPERT
Chronic severe delayed stress	4.30	3.15
Loss of arms and legs	5.13	6.12
Permanent paralysis below the neck	5.35	6.12
Regular severe pneumonia	4.65	3.85

APPENDIX J. DATA FROM THE RISK MATRICES

The tables on the following pages summarise the numerical data entered into the risk evaluation procedure, following the format given in Table 4. The hazards are listed in order of descending risk index, as shown in Tables 5 and 6, and in the risk categories shown in Table 1.

The classes of harm are as defined in Chapter 2.

In addition to the indication of orders of magnitude, a more precise estimate of number of people affected is given for domestic safety issues; this reflects the availability of better statistics.

The number of stars in a cell of the matrices reflects strength of the evidence, as defined in Chapter 3.

Table J1. First level of risk - health issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED					
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1
Hygrothermal conditions	I		*	***			
	II		**	***			
	III	*	**	***			
	IV	***					
Radon (domestic)	I			***			
	II			*			
	III						
	IV						
House dust mites	I			*	**		
	II			**			
	III	*	**	***			
	IV	*					
Environmental tobacco smoke (domestic)	I			*	**	***	
	II			*	**		
	III		***				
	IV	***					
Carbon monoxide (domestic)	I				**	***	
	II				*	**	***
	III		*				
	IV	**					
Radon (non-domestic)	I				**	***	
	II				*		
	III						
	IV						

Table J2. Second level of risk - health issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED					
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1
Environmental tobacco smoke (non-domestic)	I				*	**	
	II			*			
	III		*				
	IV	***					
Fungal growth (domestic)	I				*	**	
	II				**		
	III		*	**			
	IV	***					
Security and the effects of crime (domestic)	I					**	
	II			*	**	***	
	III	*	**	***			
	IV	***					
Sources of infection other than sanitary accommodation (non-domestic)	I				*	**	
	II				*		
	III				*		
	IV	*	***				
Noise (domestic)	I					*	**
	II				*		
	III	*	**				
	IV	***					
Carbon monoxide (non-domestic)	I					**	***
	II				*	**	
	III			*			
	IV			*			
Lead (domestic)	I						***
	II					***	
	III		***				
	IV		*				

Table J3. Third level of risk - health issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED					
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1
Sanitary accommodation (domestic)	I					*	**
	II					*	
	III		*	**	***		
	IV	*	**	***			
House dust mites (non-domestic)	I						
	II						
	III		**				
	IV	*					
Sources of infection other than sanitary accommodation (domestic)	I					*	**
	II				*		
	III			*	**		
	IV	*	***				
Space (domestic)	I						*
	II				*		
	III		*	**			
	IV	***					
Volatile organic compounds (domestic)	I					*	
	II					*	
	III				*		
	IV	*	**				
Volatile organic compounds (non-domestic)	I					*	
	II					*	
	III				*		
	IV	*	**				
Oxides of nitrogen (domestic)	I						*
	II				*	**	
	III		*	**			
	IV	*	**				
Space (non-domestic)	I						
	II						
	III			*			
	IV	***					
Lighting (non-domestic)	I						*
	II					*	
	III				*		
	IV	***					
Hygrothermal conditions (non-domestic)	I						
	II						
	III						
	IV	***					
Particulates (domestic)	I						**
	II						*
	III						*
	IV		*				
Fungal growth (non-domestic)	I						*
	II					*	**
	III				*	**	
	IV	*	**				

Table J4. Fourth level of risk - health issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED					
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1
Sulphur dioxide and smoke (domestic)	I						*
	II					*	
	III			*	**		
	IV		*				
Sanitary accommodation (non-domestic)	I						*
	II					*	
	III			*	**	***	
	IV		*	**	***		
Noise (non-domestic)	I						
	II						
	III			*			
	IV	*					
Landfill gas (domestic)	I						*
	II					*	
	III				*		
	IV			*			
Oxides of nitrogen (non-domestic)	I						
	II						
	III						
	IV		*				
Pesticides	I						
	II						
	III					*	
	IV				*		

Table J5. First level of risk - safety issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED						
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1	Estimate
Slips, trips and falls on stairs and ramps (domestic)	I				***			540
	II			***				5,100
	III		***					27,800
	IV		**					66,000
Slips, trips and falls on the level (domestic)	I				***			115
	II			***				9,300
	III			***				9,300
	IV		**					59,100
Burn and scald hazards (domestic)	I					***		80
	II		***					8,000
	III		***					32,000
	IV		***					29,000
Fuel gas (domestic)	I				***			254
	II					***		56
	III			***				3,100
	IV			***				3,300
Slips, trips and falls from buildings (domestic)	I					***		77
	II				***			690
	III				***			890
	IV			**				2,000
Slips, trips and falls on the level (non-domestic)	I				***			
	II			**				
	III		*					
	IV		*					
Drowning hazards (domestic)	I					***		72
	II					***		43
	III					**		40
	IV							0
Slips, trips and falls involving baths (domestic)	I					***		55
	II				***			540
	III			***				2,300
	IV			**				6,700

Table J6. Second level of risk - safety issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED						
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1	Estimate
Slips, trips and falls on stairs and ramps (non-domestic)	I					***		
	II			***				
	III			***				
	IV		**					
Electrical hazards (domestic)	I					***		30
	II				***			133
	III				***			855
	IV			**				950
Slips, trips and falls from buildings (non-domestic)	I					***		
	II			**				
	III			**				
	IV			*				
Architectural glass (domestic)	I					***		11
	II				***			360
	III			***				1,400
	IV			***				4,000
Electrical (non-domestic)	I					***		
	II				***			
	III			***				
	IV							
Collision/entrapment involving windows (domestic)	I						**	1
	II				***			475
	III				***			874
	IV			***				5,300
Explosions (domestic)	I					***		10
	II					***		37
	III			***				2,000
	IV			**				2,200
Explosions (non-domestic)	I					***		
	II					*		
	III				*			
	IV			*				

Table J7. Third level of risk - safety issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED						
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1	Estimate
Collision/entrapment involving doors (domestic)	I						***	3
	II						**	0
	III				***			220
	IV		***					24,100
Falling objects (domestic)	I						**	3
	II					***		16
	III				***			880
	IV				***			950
Falling objects (non-domestic)	I						***	
	II						**	
	III				*			
	IV							
Collision/entrapment involving doors (non-domestic)	I						**	
	II						**	
	III					**		
	IV			**				

Table J8. Fourth level of risk - safety issues

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED						
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1	Estimate
Collision/entrapment involving windows (non-domestic)	I							
	II					**		
	III				**			
	IV			**				
Collision/entrapment involving lifts (domestic & non-domestic)	I							
	II						**	
	III				**			
	IV				**			

Table J9. Fire hazards

HAZARD	CLASS OF HARM	NUMBER OF PEOPLE AFFECTED						
		$>10^5$	$>10^4$	$>10^3$	$>10^2$	>10	>1	Estimate
Collision/entrapment involving windows (non-domestic)	I				***			
	II				**			
	III			**				
	IV			**				
Collision/entrapment involving lifts (domestic & non-domestic)	I							
	II							
	III							
	IV							