DIMENSIONS OF PERCEPTION AND RECOGNITION OF DANGER

by: Michel Pérusse

Thesis submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy (Ph D)

The University of Aston in Birmingham

September 1980

SUMMARY

Despite abundant literature on human behaviour in the face of danger, much remains to be discovered. Some descriptive models of behaviour in the face of danger are reviewed in order to identify areas where documentation is lacking. It is argued that little is known about recognition and assessment of danger and yet, these are important aspects of cognitive processes. Speculative arguments about hazard assessment are reviewed and tested against the results of previous studies. Once hypotheses are formulated, the reasons for retaining the repertory grid as the main research instrument are outlined, and the choice of data analysis techniques is described. Whilst all samples used repertory grids, the rating scales were different between samples; therefore, an analysis is performed of the way in which rating scales were used in the various samples and of some reasons why the scales were used differently. Then, individual grids are looked into and compared between respondents within each sample; consensus grids are also discussed. The major results from all samples are then contrasted and compared.

It was hypothesized that hazard assessment would encompass three main dimensions, i.e. 'controllability', 'severity of consequences' and 'likelihood of occurrence', which would emerge in that order. The results suggest that these dimensions are but facets of two broader dimensions labelled 'scope of human intervention' and 'dangerousness'. It seems that these two dimensions encompass a number of more specific dimensions some of which can be further fragmented. Thus, hazard assessment appears to be a complex process about which much remains to be discovered. Some of the ways in which further discovery might proceed are discussed.

Michel Pérusse, Ph D thesis, 1980

Key words : COGNITIVE PROCESSES
HAZARD ASSESSMENT
REPERTORY GRID

ACKNOWLEDGEMENTS

The author wishes to acknowledge the collaboration of many persons without whom this research would not have been possible. Firstly, Mr A Champion, research officer at the Steel Casting Research and Trade Association, and Mr CH Green, researcher at the School of Architecture of the University of Dundee, kindly granted permission to use for the purpose of this thesis, data collected within their own researches; the author would like to thank them very sincerely. Secondly, Mrs E Maclean, director of home safety in the Royal Society for the Prevention of Accidents kindly granted permission to use a RoSPA drawing for research purposes; her assistance must be acknowledged with gratitude. Gratitude must also be expressed to Dr Melinek of the Fire Research Station, for his authorizing the author of this thesis to quote from an unpublished research report. Dr A I Glendon, lecturer in the Department of Safety and Hygiene at the University of Aston in Birmingham, carefully read all of this thesis at its various stages; his patience and valuable comments deserve profound gratitude. The author is also indebted to Miss Anne Lambert and Mrs Monique Boyce whose secretarial assistance in the preparation of this thesis included typing as well as valuable comments on presentation. Dr A R Hale, senior lecturer in the Department of Safety and Hygiene, supervised this research; his help and understanding were available in difficult situations not all of which were directly related to this thesis; without such help this thesis would never have been completed. Last but not least, the author's wife, Liane, typed many drafts of chapters and provided suggestions on lay-out as well as on contents; relinquishing holidays was but one form of her moral support; words are hardly sufficient to express all the gratitude which such collaboration deserves.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 : HUMAN FACTORS IN SAFETY

Choosing a research topic within the broad area of human factors in accidents is not an easy task. Firstly the terms "human factors" have to be defined. Secondly the relative importance of the various factors in accident causation needs to be assessed.

Various authors have proposed models which describe sequences of human factors. This first chapter is an analysis of these models in order to identify potential areas of research need. Researches are reviewed, in Chapter 2, which supply indications about the relative importance of some human factors.

1.1 THE NEED FOR A FRAMEWORK

It is not easy to evaluate the considerable volume of published material on the role of human factors in safety. Surry (1969) reviewed 246 publications while Hale and Hale (1972) reviewed 355 documents. Even allowing for the 57 publications reviewed in both books, these figures represent a large number of documents on the topic. Since these reviews interest in safety matters has continued to grow and many further publications have appeared.

Given such a quantity of information, in order to summarize what is known about human factors so as to show where knowledge is wanting, frameworks and models are almost essential.

In using models two strategies can be adopted. Firstly new models may be derived from a thorough and critical literature review. Alternatively existing models may be examined with a view to their improvement. Models are usually devised either by authors who have carried out a major review of the literature, or by researchers who incorporate a major development to an existing model on the basis of empirical findings. When a problem needs to be defined for further investigation comparing models can be valuable in terms of time.

1.2 THE USE OF EXISTING MODELS

A number of models have been devised to synthesize human factors in relation to danger and accidents. Some models, like Suchman's (1961) and Wigglesworth's (1972), are relatively simple and limited in scope. Others are more elaborate and tend to incorporate the simpler models. At this stage a broad overview of human factors is sought. Therefore discussion in the rest of this chapter will focus upon brief analyses of the broader and more elaborate models.

Models may be seen as representations of complex realities in summarized and simplified form. Very often they are devised for a specific purpose and therefore to critize a model in great detail can be misleading. General criticisms of a number of models, however, can serve to focus attention upon topics worthy of further consideration.

The first model to be discussed will be that of Surry (1969). Her model was the first major attempt at providing a structure for the evidence which existed at that time. Two of the other three models discussed in this chapter were elaborations upon Surry's model. Therefore Surry's model will be analysed in greater depth than later models. Later models will be reviewed mainly in terms of features which were added to Surry's model. Then similarities and differences between the four models will be reviewed.

1.3 A DECISION MODEL OF THE ACCIDENT PROCESS

Surry (1969) proposed a two-stage model as an explanation of the process by which accidents occur. The first stage, or cycle, of the model is concerned with an increase or build up of danger. The second cycle is a direct consequence of the first, and is initiated when danger has built up to the extent that it is being released. This "decision model of the accident process", discussed below, is shown in Figure 1.1.

Both cycles of the model are made up of four parallel types of components. The first category of components in each cycle is essentially environmental; it concerns the presence of perceptible indications that danger is building up or being released. The second pair of components are perceptual and refer to whether the information is perceived by a person involved in the process. The third types of components refer to cognitive processes. A series of questions are successively asked in each case: A) is the meaning of the perceived information recognized? B) is the right mode of avoidance known? and C) is a decision made to try to avoid the danger? Surry (1969) labels the fourth

FIGURE 1.1

A DECISION MODEL OF THE ACCIDENT PROCESS (SURRY, 1969)

Aston University

Illustration removed for copyright restrictions and final component type "physiological response". The question which is asked for that component in both cycles is: does the person have the physical capability of avoiding the danger?

Surry (1969) postulates that, in the first cycle, if all questions are answered affirmatively "the danger will not grow and no injury can ensue". But if any question is answered negatively "the danger will become imminent". In the second cycle, assuming that danger has built up, Surry (1969) suggests that an affirmative answer to all questions will lead to an accident being avoided. But a negative answer to any question "will lead inevitably to injury" or to damage.

1.4 DISCUSSION OF THE MODEL

1.4.1 Some assumptions

The model in Figure 1.1 makes a number of assumptions, some of which are interdependent. Firstly it is assumed that danger is present and building up. This assumption in turn has a few implications. For instance the model can be used to describe neither the place from which danger arose nor what type of danger was present. Danger simply seems to emerge from a situation called "Man and Environment" like a deus ex machina.

It is therefore also assumed that the type of danger which is present has no influence on the model. But it may be that some types of danger (either environmental hazards, human errors, mechanical hazards or others), because of their very nature, influence course of events consistently in the same direction. The model is not likely to identify such dangers easily.

Because the model is concerned neither with the origin nor with the nature of danger, it can only be used to describe how an accident occurred once danger was present rather than why an accident occurred. Thus it could not be used to identify frequent sources of danger or prescribe what should be done to alleviate them. And yet, eliminating danger at source is probably one of the most effective prevention strategies.

1.4.2 Linearity of the model

The model which Surry devised is both linear and sequential, having neither feedback nor monitoring loops. But the very nature of the phenomenon which the model attempts to describe would justify including such loops. As it stands the model conveys the impression that encounters with danger are one-off events, whereas such encounters can be continuous. For instance there are phases during which danger is building up and which warrant intervention or "avoidance" only if certain conditions are met. Initially these phases only require monitoring. Pressure build-up inside a boiler is an example of this. In such instances information is processed not once but a number of times. Some form of loop might be required in the model in order to account for those instances.

There are other types of situations in which one course of action is not enough to avoid danger completely. For instance avoiding one danger may lead a person directly into another danger. There may also be instances where a person takes an appropriate avoidance action but his efforts are nullified by someone else's actions. In these instances new circumstances make it necessary for a person to re-appraise the situation. Some form of loops might be needed in the model in order to suggest the possibility of such re-appraisals.

Surry (1969) argues that the sequence of steps in her model is re-initiated with every new situation. This may be taken to be equivalent to a feedback mechanism. Whether "re-starting" the model adequately covers all the eventualities discussed above is problematic.

1.4.3 Simultaneous dangers

The linearity of Surry's model is the cause of another ambiguity. Surry argues that the model is applied to all dangers present in a given situation. Certain dangers, however, may need more urgent consideration than others, and some dangers may be inter-dependent. If urgent, inter-dependent or both types of dangers are present, then the model may need to be applied either simultaneously, or iteratively or both to all dangers. How this can be done is not obvious from Surry's model.

In the danger buid-up cycle it is taken for granted that a person who knows what to do in order to prevent danger from building up but does not take the appropriate actions is taking a risk. However, the person may decide not to remedy the situation straight away in order to prevent a more immediate, and possibly a more dreadful, accident. In such a case, although strictly speaking the label "risk-taking" is appropriate, "logical behaviour to reduce injury" might be a more relevant description.

The way in which a person copes with simultaneous dangers does not appear to be accounted for in the model. It might be argued that some steps in the model represent complex decisional processes. It is possible that a decision not to attempt avoidance of a given danger is reached as a result of assessing all dangers present. But this implies a process of hazard assessment which Surry does not mention.

1.4.4 Applicability of the model

Some of the arguments discussed above may not be relevant if the model is seen as a structure for explaining events post-hoc. It may be that the model was devised in order to focus attention only on the danger which was directly responsible for a given accident. However, such a purpose assumes unicausality of accidents.

Surry (1969) challenges the assumption of unicausality. Moreover she writes that the model can be applied to all dangers present in a situation. This therefore leaves the arguments about the model's oversight of interactions between dangers and about its applicability solely for descriptive purposes open for debate.

There is another type of interaction which the model fails to take into consideration. When an accident occurs many persons can be present. These persons may or may not have a role to play in the course of events. Their roles may or may not be interactive or interdependent. For instance, a person may not be aware that he is about to be struck by a falling object; but that person may be pulled away from the path of the object at the last second by a witness who realised what was about to happen. Such interactions of roles do not seem to be accounted for by the model.

This oversight in turn raises another issue. When many persons are present at the scene of an accident they may or may not be victims. Despite its connotation of active involvement which need not be relevant, the word actors will hereafter be used to describe persons present at the scene of an accident who are not victims.

Surry's model does not appear to make any distinction between actors and victims. This absence of distinction has some advantages. For instance it realistically implies that responsibility for an accident, responsability for accident prevention and being a victim of an accident may be independent. Which factors, however, lead to a person being an actor rather than a victim or vice versa are not obvious from the model.

For that matter there may be other reasons, unrelated to the people involved, why some accidents cause injuries, some cause damage and others cause both. There may be factors which, irrespective of the people involved, explain severity of injuries and extent of danger. These factors and reasons are not incorporated in the model.

Again these absences represent advantages. They emphasize, for instance, that accidents resulting in injury are not the only ones worth investigating. On the other hand if the model is an attempt to describe the whole of the accident process, then the final segment, explaining the outcome, is missing. This may restrict the applicability of the model and make it less useful for people primarily concerned with reducing the severity or the number of injuries.

1.4.5 Summary of the discussion

The issues discussed in the section above suggest that Surry's model essentially describes a course of events. Its predominantly descriptive purpose means that it has limited use as a predictive model. Even within its descriptive role the model has some limitations.

Firstly the model ignores the source of danger. Secondly, even if the model can be applied to all dangers present in a situation, it is not designed to cope with interactions between dangers and people's reactions to simultaneous dangers.

Thirdly the model does not differentiate between victims and actors. Nor does it foresee the possibility of interactions between all persons involved. Finally, the model does not attempt to describe what outcome events are likely to lead to.

These criticisms must not be seen as directed at Surry's ability to design a model. From the literature she reviewed it seems that her model is an adequate synthesis of those human factors which had been examined up to that time. The oversights and shortcomings which have been pointed out appear to be areas which had been left relatively unexplored when Surry proposed her model.

Developments have taken place since that time and other models have been designed. Some of these have taken Surry's as a basis to expand from; others are based on a different approach. Some models cover only a few of the areas covered by Surry's; others cover new or different areas. Most of them, like Surry's, have shortcomings.

This chapter continues by discussing a few of these models. They have been chosen because of their broad scope and because they incorporate features which are absent from Surry's model. It is these new features of the models, rather than their shortcomings, which will be examined.

1.5 AN EXPANSION OF SURRY'S MODEL

One of the criticisms of Surry's model was that it assumed danger was present from the start, thus ignoring sources of danger.

Andersson et al (1978) argue that this causes problems of "classification of the causal patterns". The model which they propose is an attempt to overcome some of those problems. It is depicted in Figure 1.2.

Probably the most noticeable feature of the model is the resemblance of its second and third sections to Surry's model. Even within these sections however there are slight differences. In each section the steps which have been adapted from Surry's model have been reworded. Changes in meaning appear to be negligible.

Within each of the comparable sections Andersson et al (1978) have incorporated two additional steps to Surry's model. These steps ask the questions: A) "can the danger be avoided?" and B) "is there freedom of choice?". This introduces a distinction between a danger avoidable per se and a danger avoidable by a person involved in the process. The corresponding question in Surry's model referred to a person's ability to avoid more than to the avoidability of a danger. These questions remove two of the assumptions made in Surry's model: firstly that danger is always avoidable and secondly that a person confronted by danger can always decide whether or not to attempt to avoid it.

The main modification which Andersson et al (1978) make is to include a first section preceding the two sections derived from Surry's model. The first section was devised for the specific purpose of describing the presence of danger in

FIGURE 1.2

SURRY'S MODEL AS MODIFIED BY

ANDERSSON ET AL (1978)

work systems. It examines in some detail the starting point, labelled "man and environment", in Surry's model. Andersson et al (1978) mention that this section enabled them to shed some light on accidents which could not be analysed by Surry's model. This development of the model therefore seems justified from the theoretical point of view argued earlier as well as from the point of view of practicality for case studies.

The modified model still leaves some of the previously raised questions unanswered. Some of these questions are also mentioned by Andersson et al (1978). Finally, although the new model does describe the presence of situational danger, it can hardly be applied to describe danger which would be attributed to human intervention in the work system. This however constitutes the essence of the next model to be discussed.

1.6 A MODEL OF ACCIDENT CAUSATION

At about the same time as Surry (1969) proposed her model, Hale and Hale (1970) devised a model the purpose of which was to remedy some shortcomings in research on accidents. The model they proposed is shown in Figure 1.3.

This "model of accident causation" has four main steps. Firstly information is perceived by a person (whether actor or victim of an accident). This information is a function of: available information, information expected by the person, and the mechanisms influencing both these factors.

FIGURE 1.3

A MODEL OF ACCIDENT CAUSATION

(HALE AND HALE, 1970)

Secondly, given the perceived information, a range of possible courses of action are devised in order to cope with this information. The elaboration of the various courses of action is influenced by training, skills, goals, etc.

Thirdly a decision is made about which course of action to adopt. The decision is made on the basis of the person's assessment of the advantages and disadvantages of each type of action. Finally, once a course has been chosen, action is taken.

Whatever action is taken, it influences the prevailing situation. That influence gives rise to modified, different or additional available information with which the person will have to deal. This is the mechanism which is implied by the loop in the top part of Figure 1.3.

1.7 DISCUSSION OF THE MODEL

1.7.1 Feedback loop

It was argued in the discussion of Surry's model that its linearity conveyed the impression that encounters with danger were one-off events isolated from other events in a situation. Andersson et al (1978) criticize Surry's model in much the same way. The feedback loop in the Hale and Hale (1970) model is therefore one of its important features.

There is nothing in the model which specifies that information present in the situation is related specifically to danger. Information can be related to anything, including danger.

This helps to convey the notion that danger can occur both within the context of normal human actions, and as a result of it.

1.7.2 The notion of causation

Because the model implies that danger can stem from human actions, Hale and Hale (1970) say that it describes "causation". In other words the model can explain where danger comes from and what causes it. In that respect the emphasis is different from that of a model of the accident process. The latter describes how rather than why the accident occurred.

However, the model does not necessarily assume all dangers to be the result of human operation. Danger may arise from another source. For example, Andersson et al (1978) point out that danger can often arise as a result of the work process. Therefore the first section of the model proposed by Andersson et al (1978) and the whole model proposed by Hale and Hale (1970) would tend to complement each other in describing the origin of danger.

1.7.3 Some similarities

The purpose of Hale and Hale's (1970) model is not restricted to the description of danger as originating from human actions. The model describes the basic mechanisms by which a course of action is chosen. What course of action is chosen depends on prevailing conditions, but the mechanisms which lead to a decision remain the same, irrespective of prevailing conditions. Furthermore the model implies that the outcome of one action will influence the decision on the next step in a sequence of actions.

As pointed out earlier, the two sections of Surry's model describe basically identical processes; what changes from one section to the other are prevailing conditions or phases of danger. The same can be said of the corresponding sections in the Andersson et al (1978) model. Hale and Hale's model emphasises this repetition of the process by enclosing it within a loop.

Moreover, the categories of components show similarities between models. In all three models there are steps dealing with availability of information, perception of that information and decision upon alternative behaviours to be adopted. The components which refer to availability and perception of information are relatively similar throughout the models. Decision-making mechanisms on alternative behaviours, however, are not treated in the same way in all models.

Despite these differences, the importance of decision-making mechanisms is recognized in all three models. For instance, Surry (1969) calls her model "a decision model". One of the steps which Andersson et al (1978) add to Surry's model queries a person's freedom of choice. Similarly it appears from their model that Hale and Hale (1970) consider choice and decision to be rather elaborate mechanisms. Nevertheless these mechanisms result in the greatest differences between comparable sections of the three models. At first sight the next model to be discussed presents the same mechanisms in yet another way.

1.8 A STRUCTURE FOR EMPIRICAL EVIDENCE

Hale and Pérusse (1977) have discussed results of research undertaken by themselves and their colleagues. In an attempt to provide a theoretical structure for the empirical evidence they reviewed, they propose the model depicted in Figure 1.4 (Hale and Pérusse, 1978).

Like Surry's model, this structure assumes the presence of danger from the start. Although presentation is different, the elements of the model are largely inspired from that of Surry (Hale and Pérusse, 1978). The model also assumes that the process is identical at various phases of danger. No feedback loop is shown in the model; but it is assumed that questions in the model are asked again if danger is increased. The model is therefore comparable to Hale and Hale's model when the information which is being processed in the latter is related to danger.

Steps 1, 3, 4, 6 and 7 of Hale and Pérusse's model are comparable to the questions in Surry's model. But steps 4, 6 and 7, when studied carefully, represent elaboration of corresponding steps in Surry's model. For instance, recognition and labelling of danger appear to be considered as two distinct mechanisms. Whether the appropriate avoidance action is known and whether that action can be carried out are also portrayed as two separate considerations. These may well be what Surry (1969) sees as two of the steps in her model: 1) is avoidance action known? and 2) can action be taken? Finally, step 7 appears to query whether a decision is made and, if it is, whether it receives sufficient priority.

FIGURE 1.4

A DESCRIPTIVE MODEL OF HUMAN BEHAVIOUR IN THE FACE OF DANGER (HALE AND PERUSSE, 1978)

Hale and Pérusse's model also includes steps which are not obvious adaptations of earlier models. Step 2 is an example of such an addition. It may be that, even if danger is not perceived, the person confronted with danger (whether actor or victim) checks more thoroughly-possibly with the use of special instruments-whether danger is present. This behaviour is encompassed by step 2.

The model is also alerted to the possibility that a person having recognised danger, does nothing about it on grounds that such action is outside the scope of his responsabilities. Finally the model points out that there may be circumstances where avoidance action may ward off danger only temporarily or partially.

1.9 SIMILARITIES BETWEEN MODELS

Surry's model and the other two adapted from it share three stages in common: availability, perception and interpretation of information about the presence of danger. There appears to be a difference in Hale and Hale's model about perception and interpretation. But the difference is only superficial; Hale and Hale (1970) appear to include in their definition of perception both perception and interpretation as defined in the other three models.

The proposers of all four models also agree that the process which they describe in their own way is a repetitive one. This is illustrated by Hale and Hale in the form of a feedback loop. Although such a loop is not present in the drawing of their model, Andersson et al (1978) argue that it is implied and that the process is in fact a repetitive one. Surry (1969) and Hale and Pérusse (1978) have not incorporated a loop in their model. But their assertion that their respective model is "re-started" with every change in a situation can most probably be interpreted as being in agreement with the opinion expressed by Andersson et al (1978).

Andersson et al (1978) further argue that their model can be applied to "near misses" as well as to accidents. Surry (1969) makes the assumption that whether an event will lead to a "near miss", to injury or to damage depends upon chance factors. Wigglesworth (1972) explains this assumption by saying that outcomes are dependent upon a host of factors which in turn are each dependent upon a further series of factors, and so on. Many sets of circumstances can lead to the same outcome; and some sets of circumstances can have many different outcomes. It is so difficult to establish precise links between circumstances and outcomes that to all intents and purposes chance factors govern the outcome of an accident. Perhaps this is why none of the models is concerned with the tricky issue of the outcome of accidents.

Another common feature of the four models is that none of them draws a distinction between actors and victims. This absence of distinction may be associated with lack of concern about outcome. More probably it reflects the authors' intention that their model should apply to all persons involved in an accident. As pointed out earlier, Wigglesworth (1972) argues that "chance factors" will determine whether an actor will be a victim of an accident. However, if there are "human factors" which affect outcome, it would seem that these factors are overlooked by the four models. The models could nevertheless be useful in the identification of human factors potentially affecting outcome. For instance if it was found that victims consistently follow a given path within a model whereas actors (i.e. non-victims) consistently follow a different path, one could infer that some human factors do have an influence on the type of accident results.

1.10 DIFFERENCES BETWEEN MODELS

As there are features which are common to all four models, so also are there ways in which each is unique. These unique features can be classified broadly into two categories: sources of danger and cognitive processes.

1.10.1 Origin of danger

Two of the models were concerned with the origin of danger. Andersson et al (1978) have added a whole section to Surry's model. The new section traces the source of certain dangers to work processes. Hale and Hale (1970) postulate that the source of some dangers can lie in malfunctions within otherwise normal human actions.

Features about the source of danger are absent from the other two models. This is probably because their respective proposers were not concerned so much about the origin of danger as about behaviour in the face of danger - irrespective of source.

1.10.2 Recognition

Interest about behaviour is also shared by Andersson et al (1978) and by Hale and Hale (1970). All four models describe what appears to be the same process. But within that same process there are different features, most of them related to what Surry calls "cognitive processes".

The first cognitive process in Surry's model deals with recognition of a danger warning. Andersson et al (1978) refer to this process as knowledge of the warning. Hale and Hale (1970) include it with other mechanisms under the heading of perception. Hale and Pérusse (1978) imply that the process has at least two distinct mechanisms. In attempting to synthesise and integrate these notions in order to reach better understanding of this process, the resulting picture is rather elaborate. It seems that the first question to be asked is: does the warning information reach a person's senses? And if so, is it recorded by the brain? For instance, could a person give a description of the warning? These questions appear to be implied in perception (Surry, 1969) and in noticing (Hale and Pérusse, 1978).

For the warning to serve its purpose it must, in a person's mind, be associated with danger. Then it can be asked:

does the person know the warning? Has that information been seen before, heard of, experienced, processed, or whatever? In other words is the person "familiar" with the information?

Hale and Pérusse (1978) imply that understanding a warning involves more than simply knowing it. Their mention of a labelling process suggests some form of hazard assessment. In three of the four models questions are formulated to be answered by a yes or a no. Thus all questions in these models seem to refer to dichotomous or binary mechanisms. However, danger can be a matter of degree. Danger can be a function of likelihood of an accident or seriousness of a potential injury or of other such variables. Neither likelihood nor seriousness can be easily assessed as dichotomies.

In proposing their model, Hale and Pérusse (1978) argue that recognition of danger involves an assessment of the available information. It may be that such an assessment involves weighing the information either against certain standards or against other information. The outcome of the assessment may then in practice be a dichotomous decision on whether the information is a danger signal. Finally it is probable that, in the models, outcomes will influence the path to be followed by events. This line of reasoning is proposed on a speculative basis, and evidence to support it is scarce. It is, however, sufficiently coherent to explain some of the results reviewed by Hale and Pérusse (1978).

1.10.3 Identification of avoidance action

All four models attempt to explain or describe accidents. Accidents imply the presence of a degree of danger which warrants avoidance. Hence all four models include as the next stage the identification of avoidance actions.

If Hale and Pérusse's postulate is right, at this stage a person may already have decided that the warning did not signal danger. This may mean that for the person concerned the warning does not warrant action. Identification of avoidance actions is then irrelevant.

Hale and Pérusse (1978) suggest another consideration which makes thinking about avoidance redundant. They point out that there may be cases where a person does not see it as his responsibility to take appropriate action. Andersson et al. (1978) postulate that, if danger is unavoidable in the first place, any further speculation about avoidance is redundant.

All four models seem to agree on the next step. If events get so far, does a person know how to avoid danger? Hale and Hale (1970) broaden the scope of this stage by suggesting that there may be more than one possible avoidance action. Andersson et al (1978) hint in the same direction by querying whether a person is free to choose between avoidance actions. Hale and Pérusse (1978) outline one possible limitation to this freedom of choice; they point out that some avoidance actions may not be available.

1.10.4 Decision

At the next stage in all four models a decision is made by the person involved in the process. What is decided then is whether to proceed with avoidance action. Hale and Hale (1970) postulate that such a decision is made by weighing the advantages of a course of action against its disadvantages.

In addition Hale and Pérusse (1978) also postulate that the outcome of this weighing mechanism may be ambiguous. Such ambiguity might lead a person to allocate a low priority to the action to be taken. Hale and Pérusse (1978) suggest that a discrepancy between an allocated low priority and a warranted high priority can lead to an accident.

If at this stage a person decides against taking action Surry (1969) says that the person is taking a risk. Technically all risk-taking behaviours follow that path of the model. It is possible, however, that some behaviours which follow that path are not risk-taking. For instance, it may be that action is merely postponed (allocated a low priority) so as to enable a person to attend wither to a more salient or to more urgent danger. Wilful neglect of danger, usually associated with risk-taking, is not present in such behaviours.

1.10.5 Action

The next stage is concerned with the carrying out of the action if a decision is made to proceed. Strictly speaking this stage cannot be classified as a cognitive process. But differences between models at this stage raise some interesting points.

Hale and Hale (1970) and Hale and Pérusse (1978) query whether action is taken. It may be that, by the time events reach this stage, an accident has already occurred. Or it may be that appropriate actions cannot be carried out. Surry (1969) and Andersson et al. (1978) include in their model a stage which queries whether action can in fact be taken.

Hale and Hale (1970) point out that, if and when action is being carried out, differences in performance can sometimes be observed. Thus, either the same person will perform the same action a number of times in different ways, or the same action can be performed by different people in different ways.

If avoidance actions can be performed in different ways they may not all have the same result. It may be, for instance, that some avoidance behaviours will ward off danger only partially, or only temporarily. These eventualities are catered for by the last question in Hale and Pérusse's model; this stage queries whether the action is "sufficient and enough".

1.11 SUMMARY

It was hypothesised at the beginning of this chapter that an analysis of models of human factors in safety would help to outline broad areas where knowledge is wanting. Four models were reviewed and their features were compared. Some of these features were very similar in all four models. It can be argued that common features represent areas where few questions are left to answer. The next chapter shows that such an assumption is somewhat erroneous.

Certain other features appeared only in one or two of the models. In order to explain these differences, hypotheses were formulated which may need to be substantiated by empirical evidence. Some of these differences were incorporated into Surry's original model in order to account for certain research results, but research results do not account for all the differences discussed in this chapter.

Origin of danger was a component of two of the models, but was ignored in the other two. In the model proposed by Andersson et al (1978) danger is seen as stemming from the work process. To a certain extent the features of the work process whicl give rise to danger and the features of human behaviour in the face of such danger are different matters altogether. Hale and Hale's model, however, suggests that the features of human behaviour from which danger can stem and the features of human behaviour in the face of danger are essentially the same basic mechanisms. Therefore answers to questions raised in relation to these mechanisms might apply to the origin of certain dangers as well as to behaviour in the face of danger.

All four models have a stepwise structure, i.e. the outcome of one step determines whether subsequent steps apply.

Whereas the first steps are common to all models, all steps labelled as cognitive processes show differences. Thus it was decided to focus attention, in this thesis, upon cognitive processes on the grounds that they appear to involve areas where greatest uncertainty exists. Since it was pointed out in this chapter that cognitive processes are very elaborate, further delimitation of the subject will be needed. The studies reviewed in Chapter 2 suggest a specific aspect of cognitive processes where research is needed.

The steps in all four models which concern action also show differences. Whether these steps apply in given circumstances depends upon the outcome of cognitive processes. However, it may be that in practical terms cognitive processes represent less of a problem than avoidance action. Whether priority should be given to the former or to the latter depends upon their relative importance in the accident process. The substance of Chapter 2 will therefore be an assessment of the contribution made to accident causation by cognitive processes and performance of avoidance actions.

CHAPTER 2: IMPORTANCE OF COGNITIVE PROCESSES

As argued in the previous chapter, some confusion exists over the nature and role of cognitive processes in accidents. A discussion of the models was an attempt at theoretical integration of scattered fragments of evidence. In order to assess the relative importance of the various processes it was necessary to review studies which encompassed most if not all of the cognitive processes which have been discussed in relation to accidents.

Hale and Hale (1972) observe that most publications on industrial accidents are of one of two types. Some of these publications are attempts to explain accidents by matching their frequencies to various statistical distributions and their underlying assumptions. Most of the other publications on accidents describe attempts to demonstrate the relevance of some variable or other in accident causation.

It was therefore very difficult to identify pieces of research which could shed some light on the relative importance of cognitive processes. Only two such studies were located: one by Abeytunga (1979) and the other by Lawrence (1974). This chapter is essentially a review and discussion of these two studies.

2.1 A STUDY INTO SAFETY TRAINING NEEDS

The main purpose of Abeytunga's (1979) research was to identify the safety training needs of construction supervisors.

Abeytunga interviewed 70 supervisors, the interview being designed to investigate supervisors' perceptions of the hazards or "accident symptom occurrences" (ASO) which prevailed on their construction sites.

The standard procedure first involved a tour of a site by the investigator (a trained construction inspector) and a supervisor. During the tour both investigator and supervisor made separate notes of hazards they spotted. After the tour a discussion took place during which notes were compared. The investigator then asked a few questions about each hazard.

2.1.1 The interview

The schedule of questions which were used in the interview is presented as a flow chart in Figure 2.1. The schedule was applied for each ASO with each supervisor. As can be seen from the figure the schedule comprised three main stages.

Firstly it was ascertained whether a hazard on the investigator's list had been spotted by the supervisor. If it had been noted by the supervisor the second stage was carried out; if it had not been noted, the supervisor was asked whether he agreed that it was a hazard. If the supervisor disagreed the reasons for his disagreement were examined and the schedule was restarted for another ASO. But if the supervisor agreed about the hazard, then the interview proceeded to the second stage. Abeytunga (1979) does not mention any ASO noted by the supervisor but not by the investigator.

FIGURE 2.1

A FLOW-CHART OF THE INTERVIEW SCHEDULE USED BY ABEYTUNGA (1979)

Aston University

Illustration removed for copyright restrictions

Legend

A : Hazards, or Acccident Symptom Occurrences (ASO), are noted by the interviewer.

B : Did the supervisor make a note of the ASO?

C : Did the supervisor agree that it was an ASO?

D : Did the supervisor think that he should deal with the ASO?

E : Did the supervisor think that someone else should deal with the ASO?

F : Did the supervisor know of a suitable method for dealing with the ASO?

G : Was a suitable method for dealing with the ASO agreed upon?

H : Reasons sought, process terminated for one ASO and restarted for the next.

N.B. The numbers along the arrows indicate the number of ASOs discussed at each stage.

In the second stage the supervisor was asked whether he thought that he should deal with the ASO himself. If the answer was positive the interview moved on to the third stage. However if the answer was negative the supervisor had to say whether he thought that someone else ought to deal with the hazard. If in the supervisor's view the hazard did not warrant corrective action the reasons for this were investigated, the schedule was concluded for the ASO under consideration and the next ASO on the list was looked into. But if the supervisor did think that corrective action was warranted the third stage of the interview was carried out.

In the third and final stage of the interview the supervisor was asked whether he knew how to deal with the ASO. If a corrective action was known to the supervisor, he was asked why it had not been carried out. The schedule was then concluded for that ASO and restarted for the next ASO. If the supervisor could not think of a suitable method for dealing with the ASO a discussion took place during which the interviewer suggested such methods. If the supervisor did not agree with those suggestions the schedule was brought to a conclusion and the interview was resumed on another hazard. But if the supervisor did agree on any of the suggestions the assumption was made that the supervisor knew of a method for dealing with the ASO; the schedule was then continued accordingly.

2.1.2 Comparability of results

Abeytunga (1979) based his identification of safety training needs on the reasons, identified during the interviews, why hazards were overlooked or left unattended. A complete review of Abeytunga's analysis of reasons and training needs is outside the scope of the present discussion. However the various numbers presented in Figure 2.1 can serve as an indication of the relative importance of certain processes.

In Abeytunga's research a total of 659 hazards were looked into and discussed. In addition to the schedule of questions Figure 2.1 lists the number of hazards discussion of which followed the various branches of the flow chart.

In order to compare Abeytunga's results with those of Lawrence's (1974) research (discussed later in this chapter) raw numbers have to be converted into percentages. But percentages calculated as a function of the total number of hazards can be misleading. That is because the models discussed in Chapter 1 and the one used by Lawrence are mostly linear; in other words once the discussion of a hazard leaves the main line of the model it is not brought back into it later. In contrast, as can be seen in Figure 2.1 the discussion of some hazards, in Abeytunga's research, did leave the main flow of the schedule only to rejoin it later.

Therefore in order to make Abeytunga's research interpretable in terms of processes discussed earlier, percentages had to be calculated in a way which did not depend upon the total number of hazards. In the following discussion each percentage is calculated on the basis of the number of hazards whose discussion reached the stage being considered.

2.1.3 Some percentages

Of the 659 hazards noted by the interviewer while touring the various construction sites 287 did not appear on supervisors' lists. When these were pointed out to the supervisors 233 were acknowledged as hazards. In terms of the models discussed earlier it may be assumed that these 233 hazards were overlooked because of failures to perceive. It may be further assumed that the remaining 54 hazards, which were rejected even when pointed out by the interviewer, were overlooked because of failures to recognize. Should the assumptions be true, 35.3% of all the hazards noted by the interviewer were not perceived and 8.2% of them were not recognized by the respondents.

The next stage of the interview dealt with whose responsibility it was to carry out preventive action. At this stage 605 hazards were discussed. Supervisors thought that they themselves should deal with 30.1% (182) of these hazards. They saw it as someone else's responsibility to deal with a further 58.8% (356) of the hazards. They considered that the remaining 11.1% of the hazards were not worth dealing with.

Thus there were 538 hazards for which corrective action was considered. Supervisors knew of a suitable corrective action for 49.6% of those hazards. For a further 43.4% of hazards a suitable corrective action was agreed upon after discussion. No suitable action could be agreed upon for 7.1% of those hazards which were considered at that stage of the interview.

2.1.4 Discussion of results

As pointed out earlier 233 hazards were not at first spotted by supervisors but were later acknowledged to be hazards. The assumption was made that these oversights represented failures to perceive. Should the assumption be right, more than one third (or 35.3%) of all hazards were not perceived by supervisors.

Of the 287 hazards which did not appear on supervisors' lists 54 were not acknowledged to be hazards when they were pointed out by the interviewer. It was argued that in terms of the models discussed earlier supervisors had failed to recognize 8.2% of all hazards. Therefore it would appear that recognition failures were only between one fifth and one quarter of the number of perception failures.

Although based largely on Surry's (1969) model, Abeytunga's (1979) interview schedule included at least one important modification. The question about who should deal with the ASO led to interesting results. Only a small portion (11.1%) of the 605 hazards discussed at that stage were not considered worth dealing with Supervisors saw it as their duty to take action on less than a third (30.1%) of the 605 hazards. Supervisors thought that nearly twice as many (58.8%) hazards should be dealt with by someone else.

Such a difference in perceived responsibilities for dealing with hazards could indicate a tendency to off load responsibility for corrective action when a person is not directly at risk. Suggestions that this could be the case are found in the reasons given by supervisors as to why they should not deal personally with some hazards. Mentioning that a hazard was outside their scope of responsibility accounted for 12.5% of all reasons given at that stage.

Whether supervisors were right about matters of responsibility is difficult to say. Abeytunga points out that a construction site is a complex and ever changing environment; the organization of a site has to be re-adjusted at least on a daily,if not an hourly, basis. Roles, tasks and responsibilities can easily become ambiguous. Dealing with hazards is one of the areas where ambuity is likely to prevail. Such ambuity is illustrated by the fact that safety officers, site and company managers tended to disagree more than they agreed with supervisors' views on hazards (Abeytunga, 1979). Thus these and the following results must be interpreted cautiously.

Although supervisors knew of a suitable corective action for most hazards (92.9%), for only half of the hazards did they know off-hand of such action; for other hazards appropriate action had to be agreed upon after discussion. If Abeytunga's results are generalizable knowledge of preventive or corrective action might be a problem in the accident process.

2.1.5 Limitations

A number of factors contribute to limit the generalizability of these results. For instance the interview schedule focused solely on cognitive processes, and centered only upon supervisors' points of view. Furthermore whether supervisors hear of instances where the hazards discussed actually develop into accidents might radically alter their point of view on these hazards.

Abeytunga's interview schedule was largely based on models of the accident process. However the nature of Abeytunga's investigation modified these models, and the modifications thereby preclude further interpretation of the results.

Nevertheless some interesting points do emerge from Abeytunga's results. Firstly it appears that approximately one third of all hazards may be overlooked, even during a tour of inspection, by supervisors who are familiar with the environment being inspected. Secondly, rightly or wrongly supervisors considered that they themselves should deal with less than one third of hazards about which this question was raised. Finally supervisors knew off-hand of a solution to slightly less than half the hazards about which corrective action was discussed.

2.2 HUMAN ERRORS IN GOLD MINING

As mentioned earlier one piece of research was identified which looked into the importance of various aspects of the accident process. The main objective of this research was to establish priorities for an accident prevention campaign.

2.2.1 A model of human errors

A theoretical framework was needed in order to structure the findings. Therefore Lawrence decided to make use of various models of the accident process. Surry's (1969) model, described in Chapter 1, formed the backbone of Lawrence's (1974) model. But Lawrence also acknowledges influences from Goeller's (1969) and from Wigglesworth's (1972) models among others. The end product of the integration and slight modification of these various models is illustrated in Figure 2.2.

The model comprises two main sections. The upper, horizontal section of the model is concerned with the occurrence of an unplanned event as a result of the work activity and with the various possible outcomes of the unplanned event. The lower, vertical segment of the model, concerned with each participant's role, is adapted from the steps in Surry's del. Since only this second part of the model is relevant to the present discussion, it is the one mainly refered to hereafter.

Certain features of this model were derived from Surry's model. For instance it is a stepwise model, the outcome of one step determining the applicability of subsequent steps. Some questions are similarly worded in both models. A negative answer at any step (a "human error" in Lawrence's model) may contribute to the occurrence of an accident.

The model also incorporates a number of features which were not present in Surry's model. Firstly, unlike Surry's two

FIGURE 2.2

LAWRENCE'S (1974) MODEL OF HUMAN ERRORS IN ACCIDENTS.

consecutive cycles, Lawrence's model includes only one cycle which can be repeated under two different sets of circumstances. For instance the cycle is repeated if an error does not immediately result in an accident; thus it is implied that not all hazards lead to accidents even if human errors occur. The cycle is also repeated if a "secondary warning" is issued by one participant to another; the underlying assumption is that more than one person can be involved in an accident.

The presence of loops in the model also implies that the two cycles of danger build-up and danger release are in fact the same process repeated at different stages of events. It was argued in the previous chapter that this assumption is also inherent to Hale and Hale's (1970) and to Hale and Pérusse's (1978) models. Furthermore because of its loops Lawrence's model does convey the notion that hazards can be on-going rather than one-off events.

Secondly Lawrence has modified some steps of the process. For instance Lawrence considers the issuing of a "secondary warning" as different from taking direct action. Furthermore he does not make the assumption that a secondary warning or a direct action necessarily prevents an accident.

As pointed out earlier, these modifications allow for the possibility of either one or more persons being involved in the same accident process. Another assumption underlying these modifications is that a person can react in many different ways in the face of a certain danger.

Finally Lawrence inserts a new step in the sequence. The question asked at that step is whether the person concerned assessed the risk (i.e. the probabilities) of an accident correctly.

2.2.2 The research procedure

Using his model as a framework, Lawrence (1974) compiled information on 405 fatal accidents in gold mines in South Africa. Fatal accidents were chosen because "among the best documented accidents in gold mining are those that resulted in fatality and were subjected to legal inquiry". The written reports of those accident inquiries yielded information on 575 persons involved. When the reports were analysed in detail Lawrence found that, in terms of the model, 794 "human errors" had taken place.

Before the findings are discussed some precisions are needed. Firstly, not all errors led to accidents; otherwise 794 accidents would be discussed instead of 405. Secondly, not all accidents involved human error in the usual sense of the word "error"; for instance in three accidents there appears to have been no primary warning. Thirdly, although most accidents involved human error, it does not necessarily follow that the error or errors were the main cause of the accidents. Therefore it would appear that the word "error" was used for want of a better word to describe the various pieces of information gathered from the enquiry records.

Whether the victims of these fatal accidents would have labelled their behaviours as "errors" is less than certain. Although Lawrence's model implies that some "errors result immediately in accidents", it might be unjustified to associate the concept of error with that of guilt or responsibility.

The difference between the number of participants and the number of errors arises from those cases where two or more errors originated from the same participant. In most cases multiple errors were associated with underestimations of the likelihood of an accident. For instance someone who underestimated a hazard but nevertheless attempted an avoidance action which turned out to be inadequate was deemed to have committed two errors. Two errors were also counted when an underestimation led to a failure to take action. Finally in some cases participants attempted more than one avoidance mode and possibly issued one or more secondary warnings as well; each inadequate course of action and each inappropriate warning was counted as one error.

2.2.3 Some results

A breakdown into the various categories defined by the model of the 794 errors provides some valuable information about the relative importance of the different steps of the accident process.

Firstly Lawrence mentions that more than a third (36%) of all errors were "failures to perceive a warning". As for "failures to recognize a perceived warning", they accounted for 4.2%

of all errors. Approximately a quarter (24.7%) of errors fell into the category "underestimations of a hazard".

Underestimations and failure to perceive proved to be the two most important categories of errors. The next step in the model, i.e. "failures to respond to a recognized warning", was the third most important category; it accounted for 17.5% of the total number of errors. But, as Lawrence points out, 75% of these failures to respond were the result of underestimations. It was mentioned earlier that such inter-dependent errors accounted for the difference between the number of participants and the number of "errors".

When persons involved in accidents did take action in the face of danger, Lawrence's assumption was that, as the accidents nevertheless occurred, by definition their actions were inadequate. Hale and Pérusse (1978) point out that such an assumption may be unjustified since sometimes even the most appropriate actions may not be enough to prevent an accident. Nevertheless Lawrence mentions that inappropriate actions accounted for 13.7% of all errors.

A total of 71 secondary warnings were identified in the various accident reports studied. These were defined as warnings issued by some participants to others. Some 31 of these secondary warnings were judged inappropriate. These constituted the remaining 3.9% of errors.

2.2.4 Limitation of the results

As Lawrence (1974) points out, these results can be used at best only as indications. There are a number of limitations which preclude a general interpretation of the findings.

For instance, no information was available to Lawrence "about the events that are correctly handled by the workers" (Lawrence, 1974). Neither was there any information "about accidents that could, but did not, result in injury " (Lawrence, 1974).

Furthermore the accident reports which were analysed were slightly untypical. Firstly all accidents were fatal; it is difficult to say whether fatal and non-fatal accidents follow the same paths in the same proportions within the accident process. Secondly the types of accidents which were studied were not representative of the types of accidents in the gold mining industry. In addition the numbers of casualties in each type of accidents were not an accurate reflection of the existing statistics for this industry.

Finally Lawrence mentions that accident reports had not been written in terms of his models. Therefore in some instances assumptions had to be made on the basis of available evidence. Although precautions were taken to prevent examiner bias, the accuracy of such assumptions could not be verified. Nevertheless Lawrence's results yield some useful information on the accident process.

2.3 COMPARISONS BETWEEN THE TWO STUDIES

The two pieces of research discussed so far are different in many ways. For instance Abeytunga (1979) studied hazards; Lawrence (1975) analysed fatal accidents. Abeytunga was concerned mainly with cognitive processes; Lawrence dealt with the various steps of human behaviour in the face of danger.

Despite these differences, however, there are interesting similarities between the two series of results. For example in both studies 35 to 36% of hazards were not perceived. Also in both studies failures to recognize were far fewer than were failures to perceive; only 4.2% of hazards were not recognized in Lawrence's research, whereas 8.2% of hazards were not acknowledged as such in Abeytunga's interviews.

From that point onwards both research models have little in common. Therefore it would be somewhat meaningless to compare results across studies. However both researchers highlighted areas of the accident process which might deserve further attention.

For instance it was noted that supervisors interviewed by Abeytunga considered less than a third of hazards were worth them taking action on. Abeytunga also noticed that off-hand knowledge of corrective action was lacking for about half of the hazards for which corrective action was discussed. Finally, according to Lawrence one area which deserves attention and research is assessment of hazards; it accounted for nearly a quarter of identifiable errors in fatal accidents in gold mining and was directly responsible for at least a further 13% of the total number of errors.

At first sight perception and assessment of hazards, allocation of responsibility for action and knowledge of corrective action are all crucial areas which warrant closer investigation. However a closer examination of both pieces of research reviewed in this chapter narrows the order of priorities down to only one of those areas.

For instance, although both researchers have identified failures to perceive hazards as problem areas, Lawrence (1974) is able to quote a series of relatively well defined causes (e.g. "inadequate inspection technique", "obstruction to line of sight", etc.) for these failures. As for lack of knowledge of appropriate corrective action, the problem may not be as sizeable as it appears to be at first. For example, if Lawrence's results can be used as guidelines or as gauges of magnitude, inappropriate secondary warnings and inappropriate direct actions combined accounted for only 17.6% of all errors. These figures do not suggest as important a problem as some of the other figures do. Furthermore both Lawrence (1974) and Abeytunga (1979) seem to agree that a great deal of the solution to this problem lies in safety training.

Because it was not part of Lawrence's model, it is difficult to say to what extent attribution of responsibility is a problem. But it has recently become a well-documented research topic. Review articles, such as the one by Schroeder and Linder (1976), give indications that a fair number of research projects are being carried out in this field.

In comparison, Lawrence (1974) mentions that nothing is known of the causes of hazard underestimation. Therefore, of all the topics identified in this chapter, "the problem of underestimation of risk" appears to be the one which warrants "particular attention" (Lawrence, 1974).

2.4 CONCLUSION

As mentioned earlier, the results quoted in this chapter should only be regarded as suggestive of the importance of certain factors in the accident process. Various methodological considerations have been pointed out which limit the applicability of these results.

Moreover both sets of results may include a certain amount of bias. For instance, it was pointed out earlier that, in Lawrence's (1974) research, "since the accident reports had not been compiled in terms of the model, it was necessary in many cases to infer the details of human errors from the evidence available in the report" (Lawrence, 1974).

Some bias almost certainly existed in Abeytunga's results, albeit of a different type. Because the study was concerned with the reasons for the prevalence of hazards on construction sites, it is difficult to imagine that none of the supervisors felt either threatened or as if their behaviour was under suspicion. Hale and Hale (1972) point out that a similar phenomenon occured in a research by Jarry et al. (1962).

However both studies supply some concordant indications. For instance, in so far as both sets of results could be compared, percentages of perception and recognition failures were very similar in both studies. In a sense it was fortunate that the only two relevant pieces of research tended to lend credibility to each other.

Four aspects of the accident process were identified which appeared prominent. These were: perception and assessment of hazards, attribution of responsibility and knowledge of corrective action. A closer examination of these four aspects revealed that, because little is known about it, one of the four deserved more immediate attention. Hence it was decided that this thesis should be focused on the mechanisms of hazard assessment.

CHAPTER 3: HAZARD ASSESSMENT: THEORITICAL CONSIDERATIONS

It was argued in the previous chapter that little is known about the mechanisms of hazard assessment. In fact a literature review identified very few quantitative studies on this topic. Nevertheless much speculative discussion on the subject was also encountered.

A closer examination of these theoretical arguments can be summarized in two broad observations. Firstly the majority of arguments about hazard assessment mechanisms are based on analogies derived from other fields of knowledge. Secondly these arguments never lead to a precise definition of the mechanisms. Therefore hazard assessment appears to be a rather nebulous process greatly in need of clarification.

This chapter therefore reviews theoretical positions on hazard assessment and the topics in other fields of knowledge from which they were derived. In Chapter 4 these positions are tested against available empirical evidence.

3.1 DOUBLE MEANING OF RECOGNITION

In both cycles of her model Surry (1969) inserts the question: "Is the meaning of the warning recognized?" At first sight the question appears relatively straightforward. However careful consideration of this question reveals some ambiguity in the wording.

For instance Surry's question can elicit a negative answer for at least two reasons. Firstly a person may perceive a warning but may not know that a hazard is being heralded. Alternatively a person may acknowledge the presence of a hazard but does not personally consider it to be dangerous and consequently does not treat it as a hazard (e.g. a miner considering a crack in a tunnel ceiling as unlikely to lead to a roof fall). In both cases it may be said that the person does not "recognize" the hazard.

The first reason for not recognizing a hazard is a relatively straightforward one of knowledge, essentially involving a comparison of new information with a set of recognizable or familiar data. That is probably why Anderson et al (1978) have reworded Surry's question into "Does the person know the warning?" in the first cycle of their model, while in the second cycle the question now reads: "Does the person know that danger has been triggered?".

The second reason for not recognizing a hazard is essentially based upon decision-making, and is a process which may only be activated if a person "knows" the warning or the danger in the first place. This is probably the hazard assessment phenomenon discussed by Lawrence (1974).

Suchman (1961) was among the first authors to propose a model for the "analysis of the accident phenomenon" (Surry, 1969). In his model he incorporates a stage referred to as "appraisal of hazard". The results of this appraisal, he suggests, then become one of the factors taken into consideration when a decision is made whether to take a risk. Therefore it seems that Suchman is referring to both types of recognition described above. Yet Surry (1969), whose model was based partly on Suchman's, overlooks the distinction between the two types of recognition.

Hale and Pérusse (1978), however, reintroduce this distinction in their model. Their rewording of Surry's corresponding question reads: "Is it (the hazard) recognized and labelled as a danger?". Hale and Pérusse associate the notion of recognition to that of knowledge. The decision-making process in their model is encompassed by the notion of labelling. Thus their model implies two distinct decision-making processes: one dealing with the dangerousness of a hazard and the other with whether avoidance action should be taken. This distinction deserves further consideration.

3.2 DUAL DECISION MAKING

The next two steps in both cycles of Surry's (1969) model refer respectively to knowledge of avoidance action and decision to attempt avoidance. It can be argued that decision to attempt avoidance implies and encompasses the decision-making aspect of recognition. But if decision-making is one form of recognition, the decision to attempt avoidance should precede rather than follow consideration of avoidance modes. It is also possible that circumstances and events may dictate which decisions have to be taken in what order. In any case there is disagreement among authors about the sequence in which cognitive processes are activated.

For instance in terms of Lawrence's (1974) model assessment of a hazard is carried out before attention is paid to avoidance modes. In Abeytunga's (1978) interview schedule questions about who (if anybody) should deal with hazards preceded questions about corrective action; to a certain extent Abeytunga assumes that by the time corrective action is considered some form of decision (e.g. whether a hazard is worth dealing with) has already been made by the person concerned. Abeytunga mentions that ASOs which were not acknowledged to be hazards tended to be considered as not worth dealing with; furthermore the reasons mentioned by supervisors to justify both points of view were very similar. Thus both Lawrence and Abeytunga suggest that recognition (variable A in the diagram underneath) is followed by decision (variable B) which, if necessary, is followed by consideration of avoidance modes (variable C). The diagram A ⟶ B ⟶ C illustrates the sequence of steps.

On the other hand Surry (1969), Hale and Hale (1970) and Anderson et al (1978), in the models they propose, all suggest a decision-making process which theoretically occurs after consideration of the various courses of action has taken place. Furthermore Hale and Hale (1970) suggest that such a decision is of the cost/benefit type. Using the same notation as above, the diagram would then be A⟶C⟶B.

This last suggestion is somewhat different from the decision-making process described earlier as a form of recognition. Thus there could well be at least two decision-making stages involved in the accident process. These two stages are implied in Hale and Pérusse's (1978) model where labelling (i.e. recognition) is distinct from decision on action; the former stage is placed before and the latter is placed after the stage of consideration of the various courses of action. The description of this sequence in a diagram would then be: A⟶B⟶C⟶B', where both B and B' are decision-making processes.

As pointed out earlier, Suchman (1961) further suggests that both decision-making stages are inter-related, the outcome of the first being fed into the second. Lawrence's results seem to support the hypothesis that the outcome of hazard assessment has an influence on the decision to take action. In this author's research most underestimations of hazards led to failures to take action, and most failures to take action originated from hazard underestimations. Furthermore it was pointed out earlier that construction supervisors in Abeytunga's research seemed to consider the questions: "Is this a hazard?" and "Should anyone deal with it?" as synonymous.

It does not automatically follow that underestimation and failure to take action are but two facets of the same mechanisms. In Lawrence's study action was sometimes taken despite an underestimation of danger, and some failures to act occurred despite correct estimations.

3.3 THE HEALTH BELIEF MODEL

Theoretical support for the suggestion that there are distinct assessment stages can be found from another field. Becker and Maiman (1975) devised a model in an attempt to describe and predict the factors which influence the likelihood of a person's taking action to prevent illness or disease. This model, the Health Belief Model (HBM), is depicted in Figure 3.1.

The HBM has been revised by Becker et al (1977). But even after revision the model still describes assessment of a disease and assessment of the relevant preventive actions as two separate entities. Which, if any, of the preventive actions is carried out depends upon the outcome of both assessments.

In so far as the risk of a disease and the risk of an accident are comparable, the proponents of the HBM seem to agree that risk assessment is a cognitive process in its own right. In order to assess a risk a person has to know that it can be a risk in the first place. But there is probably no point in a person working out avoidance procedures for a risk which that person considers to be either non-existent or very small.

FIGURE 3.1

BECKER AND MAIMAN'S (1975) HEALTH BELIEF MODEL

There seem to be sufficient theoretical grounds for considering risk assessment as distinct from other closely related cognitive mechanisms. A brief review of literature on risk-taking reveals further details of both decision mechanisms discussed so far.

3.4 RISK-TAKING

In his foreword to Surry's (1969) book, Bales (1969) implies that all accidents which do not result from a breach of safety laws can be attributed to risk-taking. Although not in such an extreme form this view of the importance of risk-taking in accident causation is widely held. For instance Hale and Pérusse (1977) refer to evidence submitted by the Soap, Candle and Edible Fat Trades Employers Federation to the Robens Committee (1970-1972), and also quote from the Health and Safety at Work etc. Act (1974) of Great Britain. The view expressed in these sources is that risk-taking is an important problem in industry. This could be the reason why that aspect of behaviour is quite well documented (e.g. Wallach and Kogan, 1967; Hale and Hale, 1972).

Two main aspects of risk-taking can be identified in the literature on the subject: the mechanism for deciding on the level of risk within a situation, and the mechanism for deciding on a course of action. These two mechanisms are very similar to the two levels of decision discussed earlier. Therefore an examination of these two aspects of risk-taking provides some information about the nature of hazard assessment.

3.4.1 Choosing a course of action

The four models discussed in Chapter 1 had at least one common characteristic: they all included a step concerned with a decision about action to be taken. There were differences, however, about the nature of this decision process. For instance, Hale and Hale (1970) suggest that an individual chooses one among a number of possible actions. Hale and Pérusse (1978), however, suggest that the decision which is taken is whether the avoidance or preventive action should be carried out. The same type of decision is implied in the HBM presented in this chapter.

Yet another type of decision on action is suggested by Ross (1974). In her discussion of dangerous sports she mentions that a decision is made on whether high-risk activities (e.g. rock-climbing) should be undertaken.

Thus it appears that the decision-making process is very complex. It seems likely that more than one type of decision is involved in choosing a course of action (e.g. choosing between dangerous and safe activities, deciding whether to take preventive action). It is also possible that many activities can be chosen and arranged in sequence (e.g. taking precautions before undertaking a dangerous activity).

Hale and Hale (1972) mention quite a few variables (e.g. subjective estimate of risk, group norms, training, experience) which can influence the choice of a course of action.

Merz (1967) further suggests that the choice of behaviour in the face of danger is also influences by a person's perception of his own skills. Aldridge (1976) summarises Maslow's (1954) theory of needs and motivations in order to illustrate the way in which a decision is reached when a person is confronted by danger. A brief look at the decision mechanism is needed in order to understand how the variables mentioned above can influence the choice of a course of action.

Robaye (1963) discusses this decision-making process and proposes a description of some aspects of the mechanism. She argues that a person attributes a "valence" (a subjective value) to a desired goal, to the various behaviours which can lead to the attainment of the goal, and to the possible effects of not attaining the goal (e.g. accidents). These valences, along with the probabilities of attainment for each behaviour considered, are weighed one against the other; the behaviour with the greatest resulting valence will then be chosen.

The HBM discussed earlier includes a suggestion about the way in which the valence for each behaviour is attributed prior to the comparison which will result in a decision. In the model it is implied that the advantages of a preventive action are weighed up against the barriers (presumably the disadvantages) to this action. Hale and Hale's (1970) suggestion that the costs and the benefits of each behaviour are considered in order to reach a decision is very similar.

It then seems likely that motivation and other subjective variables can make certain courses of action appear more rewarding than others. Skill presumably reduces the perceived efforts (costs) inherent to certain courses of action.

It was argued earlier that, although they are distinct mechanisms, decision on hazardousness and decision on behaviour are closely related. Robaye's (1963) comments shed some light on the nature of the relationship between the two mechanisms. She argues that, when a person is confronted by the risk of an accident, the potential consequences and possibly the increase in likelihood of occurrence of the accident are sources of negative valence for certain behaviours. On the other hand, avoiding the accident or its consequences, or reducing its likelihood of occurrence give positive valence to certain courses of action.

There are cases where this assumption does not hold true. For instance, it is argued in the literature on stress-seeking (Klausner, 1968) and on motivation in dangerous sports (Klausner, 1968; Ross, 1974) that danger itself can be a source of positive valence. However, regardless of the assumption of positive or negative valence for danger, it is generally agreed that no valence is affixed to something considered as non-existent. Furthermore, Robaye (1963) argues that valence affixed to danger is proportional to the perceived level of danger. Thus the decision on hazardousness influences the decision on behaviour.

Nevertheless it remains that, as argued earlier, the two types of decision-making processes are distinct mechanisms. As pointed out earlier, Hale and Hale (1972) mention that many variables influence the choice of behaviour; Robaye (1963) argues that perceived hazardousness is but one of these variables.

There are indications that the outcome of hazard assessment may be the most important of these variables. For instance, Sell (1964) argues that underestimation of risk is very likely to lead to accidents. This view is supported by Lawrence (1974); his results indicate that many failures to take action were the direct results of risk underestimation. It was pointed out earlier that underestimations accounted for 24.6% of all "human errors" in Lawrence's research. This is probably why hazard assessment is the other aspect of risk-taking to which much literature is devoted.

3.4.2 Deciding on dangerousness

Many of the arguments put forward to explain risk-taking rest on two basic assumptions. First, it is generally assumed that, apart from pathological tendencies and from enjoyment of dangerous sports, danger is an unwanted or undesired aspect of certain activities and receives negative valence. Secondly it is assumed that, if a dangerous course of action is chosen despite the danger involved, it is because the level of danger was incorrectly assessed in the first place. These assumptions are found, for instance, in the works of Cohen and Christensen (1970), Wallach and Kogan (1961, 1967) and others.

Among the literature based on these assumptions, four main topics can be identified: perceived range of possible negative outcomes of certain courses of action, perceived likelihood and perceived desirability (or undesirability) of these outcomes, and the influence of certain variables on these three types of perception. Because in this thesis attention is focused upon hazard assessment these topics are looked into in the remaining sections of this chapter. Before this detailed discussion, however, a few comments on safety training and safety propaganda indicate the relevance of hazard assessment to certain strategies aimed at accident reduction.

3.4.3 Safety training and propaganda

As pointed out earlier hazard assessment is considered to be one of the main factors in attitudes towards risk (Robaye, 1963; Hale and Hale, 1.70). A point of view widely held is that the results of hazard assessment exert considerable influence upon the choice of a mode of action (Lawrence, 1974).

For instance, this belief is implicit in most safety propaganda (Hale and Hale, 1972). Such propaganda is aimed at changing a person's assessment of a hazard. The underlying assumption is that, by pointing out either the likelihood or the possible dreadfulness of a hazard, people will change their assessment of the hazard; this modified assessment should then lead to a change in behaviour.

Similar assumptions underlie some forms of safety training. When the hazards of a job and the severity of their outcomes are described to trainees it is assumed that the trainees' behaviour will be influenced in the way described above (Hale and Hale, 1972; Aldridge, 1976).

Some of the main hypotheses in the research on "the hidden benefits of first aid training" (Miller and Agnew, 1973; Atherley et al , 1973; McKenna, 1978) rest on those same assumptions. It is postulated that first aid training makes people aware of the range and of the severity of injuries which can result from certain situations. It is also postulated that this awareness alters peoples' assessment of those situations and thus their behaviour. These arguments are discussed again later in this chapter.

The aim of safety training and safety propaganda is to influence behaviour in the face of danger. It is assumed that this influence can be achieved by a change in hazard assessment. The usefulness of efforts to modify hazard assessment may well depend on what is known about this mechanism.

3.5 ASSESSMENT OF PROBABILITIES

Assessment of probabilities is the aspect of hazard assessment which received greater attention. Some of the arguments proposed in order to describe probability assessment have been developed from research in related fields (e.g. gambling). Other arguments stem from the study of the influence of certain variables on probability assessment. All these arguments are viewed in this section.

3.5.1 Gambling

One topic which may shed some light upon assessment of probabilities is the psychology of gambling. Much research has been carried out on this subject (see for example: Cohen, 1964, 1970 b ; Cohen and Chesnick, 1970; Cohen and Christensen, 1970; Cohen et al., 1969; Lichtenstein, 1965; Slovic, 1962; Edwards and Slovic, 1964; Edwards et al , 1965).

The nature of gambling has been said to be similar to that of risk-taking (Cohen, 1964). The way in which people go about gambling and the way in which they choose a course of action in the face of danger are both influenced by probability assessment. Thus knowledge about this aspect of gambling may prove useful for the comprehension of hazard assessment.

One important finding has emerged from studies of gambling which appears to have direct relevance to the study of hazard assessment. Some evidence suggests that, when a person assesses the probabilities of certain outcomes, mathematical probabilities are not the only (and possibly not even the main) criterion used.

For instance Cohen (1964) refers to situations the outcomes of which are all equally likely. Although the outcomes are totally independent of skill or manipulation, Cohen observes that people tend to express a marked preference which is often based upon the way in which outcomes are presented. For example Cohen states that, if asked whether they prefer either a 1 in 10 chance or a 10 in 100 chance, people more often choose the latter on the grounds that it gives them more chances of winning.

Other variables, such as the "subjectively expected utility" (Lichtenstein, 1965) and monetary cost (Cohen, 1964), are taken into consideration, sometimes interactively and in a mutually dependent way (Cohen, 1964 1970 b; Lichtenstein, 1965; Wallach and Kogan, 1967) with mathematical probabilities. That is why, as Lichtenstein (1965) and Wallach and Kogan (1967) point out, most studies on gambling have moved their focus towards what Cohen and Chesnick (1970) call "psychological chances". But that does not mean to say that mathematical probabilities are left out of the reckoning altogether (Cohen, 1964).

3.5.2 Influences on probability assessment

The variables mentioned above are taken into consideration along with mathematical probabilities. However, other variables have been identified which may alter, bias or override assessment of probabilities.

For instance there is some evidence to suggest that belief in one's luck or skill distorts the assessed probabilities. Some individuals believe that their luck or their skill can actually increase the likelihood of desired outcomes, in some instances even if the probabilities are heavily against those desired outcomes occurring (Cohen, 1970 a; Cohen and Christensen, 1970). This type of belief is discussed in greater detail later in this chapter.

Another variable which alters assessed probabilities is performance under the influence of alcohol. Cohen et al (1958) mention that bus drivers under the influence of alcohol grossly overestimate their chances of driving their vehicle through a gap without it touching the side posts.

There are a number of other such variables which influence the outcome of probability assessment. But these variables are not related to gambling; therefore they are discussed in subsequent sections of this chapter.

3.5.3 Limitations upon inferences from gambling studies

There are limits to the inferences which can be drawn about the nature of hazard assessment from the literature on gambling. The limits are imposed by basic differences between gambling and facing danger. Ross (1974) summarises these differences in this way:

> "Laboratory experiments on risk-taking tell us little about the danger motive in real life. There are many different kinds of risks, and people do not respond in the same way to all of them. Laboratory experiments are usually concerned with gambling and decision-making, where the probabilities and the financial rewards can be specified exactly. In real life, neither the probabilities, nor the value of the rewards and penalties, can be measured precisely. The rewards and penalties are also different in kind: the subject in a psychological experiment stands to gain nothing but a little money and to lose nothing but the possibility of more money; in real life he may stand to win or lose a fortune, and he may risk death for the sake of glory."

Ross's (1974) arguments encompass both types of decisions (on danger and on course of action) described earlier. She also raises the issues of probabilities and of nature of outcomes. For instance, probabilities associated with gambling and probabilities associated with hazards are not of the same order of magnitude. The working party on Acceptability of Risks (1977) say of gambling that 'a horse at 100:1 is a "long shot'". But, as they point out, "this is quite different from estimating risks (i.e. hazards) whose odds lie between 10,000,000:1 and 10,000:1". What the working party are implying here is that, while gambling introduces a number of biases into the assessment of odds, this is made even more unrealistic in hazard assessment by the sheer mangitude of probabilities involved.

Furthermore the very nature of possible outcomes is likely to account for important differences between gambling and hazard assessment. The outcomes which confront a gambler are discrete and limited in range; in other words a throw of dice, a spin of a roulette wheel or a draw in a lottery have but one outcome, which is one of a finite number of mutually exclusive outcomes. In contrast an accident can be one of many simultaneous outcomes of the same event; for example an operator can complete successfully a sequence of actions and bump his shin on a guard rail while doing so. And, in turn, accidents can lead to a number of different, and possibly simultaneous consequences. Whereas a gambler usually knows the full range of possible outcomes of a gamble, a potential victim may not even be aware of the possibility of an accident when confronted by a hazard.

It can also be argued that gambling and facing danger entail two distinct learning processes. If wins and losses in gambling, and succumbing to and escaping from danger are viewed as behaviouristic reinforcements, then the reinforcement schedule in gambling is quite different from that in facing danger. Kimble (1968) points out that positive reinforcement (e.g. wins in gambling) is used to "shape" behaviour, in other words to teach a new behaviour or to maintain or increase an already acquired behaviour. In contrast negative reinforcements (e.g. accidents) are used to "extinguish" behaviour, in other words to eradicate or diminish an already existent behaviour. Whereas positive reinforcement creates contentment, negative reinforcement usually generates frustration. The psychological processes activated by both schedules are thus quite different.

These considerations illustrate the reasons why it is difficult to transpose findings from the literature on gambling to the analysis of the mechanisms of hazard assessment. However some of the notions which have been highlighted by studies on gambling do seem to find confirmation in the research on hazards.

3.5.4 Probabilities of hazards

In a study of attitudes towards risk Melinek et al (1973) asked questionnaire respondents to rank a number of hazards in order of likelihood of occurrence to them personally. The researchers compared the rankings with objective estimations of probabilities. The results showed that, although some respondents could rank the hazards accurately, in a majority of cases there were discrepancies between subjective and calculated probabilities. As pointed out earlier Sell (1964) and Lawrence (1974) argue that such discrepancies can have unpleasant consequences.

One of the hazards which appears to have caused discrepancies in the research by Melinek et al (1973) was "being in your home when it catches fire". The authors state that the ranking of this particular hazard is biased by some respondents' beliefs in their own knowledge, skill and experience. For instanc some respondents who were trained firemen said that they would know exactly what to do if their home caught fire; thus their confidence in their ability to cope led them to underestimating the probability of occurrence of the hazard.

This observation may be interpreted as evidence that hazard assessment is influenced by avoidance assessment. It certainly does suggest that, although assessment of probabilities may intervene, it is not the only factor taken into consideration when hazards are assessed. If mangitudes of odds quoted by the working party on Acceptability of Risks (1977) are typical, probability assessment may in fact prove to be only a minor factor and one which is influenced by other factors. For instance, Ross (1974) suggests that perception of prevailing environmental conditions (e.g. "present snow conditions" for climbers) and confidence in one's equipment have an important bearing on the result of probability assessment.

There is also evidence which suggests that assessment of probabilities is influenced by first aid knowledge. McKenna (1978) found that certificated first aiders tended to increase their assessment of the likelihood of an accident.

Another intervening variable is suggested by Ross (1974). She writes that a person's "estimate will probably suffer from "small sample bias" ans be strongly influenced by the most recent events". For instance, if a climber "knows that several competent climbers were recently killed in an avalanche, he may over-estimate his own chances of such a death" (Ross, 1974).

3.5.5 Probabilities and learning

Ross (1974) also suggests that "repeated exposure to danger does not necessarily lead to a true appreciation of the hazards". Some evidence tends to suggest that exposure does influence assessment; but some ambiguity exists over the way in which influence is exerted.

Hale and Hale (1972) point out that accident repeaters tend to rate the risks of their job higher than do people who have not suffered accidents. One hypothesis put forward to explain such an observation is that experiencing accidents increases perceived likelihood of those accidents. Support for such an hypothesis is found in Hale's (1971) observation that some operators who had never succumbed to dangers noticed by an observer did not consider those dangers worth mentioning; those operators who had suffered consequences arising from the same dangers did mention them. Further support for the hypothesis is found in a study by Robaye et al (1963). The authors state that accident repeaters tended to overestimate the likelihood of an accident.

There is also evidence which suggests that exposure decreases perceived likelihood. For instance, Ross (1974) mentions that experienced climbers rate the hazards of a given climb lower than do less experienced climbers. It would seem that learning processes are involved in hazard assessment; but how these processes operate is unclear.

People also learn about dangers from seeing other people succumb to those dangers. "They can...learn vicariously: the successes and failures of others teach them what is possible" (Ross, 1974). "In other words operators learn to appreciate risks by profiting from their own mistakes and those of their neighbours" (Hale, 1971).

3.5.6 Margin of safety

Another concept is sometimes associated with probability assessment. It is the notion of "margin of safety". Hale and Hale

(1972) argue that a margin of safety "suggests...that people regulate their behaviour to allow for a contingency factor between their actions and the behaviour that they perceive as dangerous". In other words, it seems that people decide on a level of probability which they are not prepared to exceed. They then decide on a course of action involving a probability of accident lower than the maximum level to allow for a possible error in their assessment of probabilities.

Dunn (1971) quotes arguments which give a different definition of the margin of safety. According to him, it is argued that people seek to maintain the perceived probability of an accident constant within a chosen course of action. For instance, "in mountain climbing, as the climber increases his climbing skill, or obtains new safer equipment, he will turn to progressively more difficult climbs to maintain this margin of safety".

The first of these descriptions implies that there is a maximum probability of accident beyond which people are not prepared to undertake an activity. In contrast, the second description implies that there is also a minimum probability below which an activity would not be considered to be challenging. Thus, the word "margin" appears to be very appropriate.

3.5.7 Probabilities and control

Firemen in Melinek at al's (1973) study stated that they could cope adequately in case of a fire in their home. This confidence is their ability to cope with consequences could lead to hazard under-estimation. As pointed out earlier this suggests that assessment of outcomes is taken into consideration along with probability assessment. It also suggests that confidence in coping ability in the event of occurrence may become generalized into confidence in ability to prevent. This may be why Ross (1974) states that confidence in one's own skill, knowledge and experience intervenes in hazard assessment.

Both types of confidence described above appear to be conveyed in the notion of "control over danger" which has recently emerged from the literature on hazards. The perceived degree of control over events is also a function of a person's confidence in his own skill, experience and equipment and of his perception of prevailing conditions (Ross, 1974).

Perceived control, in turn, appears to play an important role in hazard assessment. "There are some events over which we have almost no control...Predictable disasters are less dangerous because they can normally be avoided" (Ross, 1974). Kates (1967), Wilson (1975), Green and Brown (1976), Williams (1976), and Hale and Pérusse (1977, 1978) also stress the importance of perceived control.

That there is a close link between perceived control and subjective probabilities is suggested by Cohen (1964). He mentions that people who gamble a lot believe that the more they gamble the better they are able to predict the outcome of a gamble; this supposedly improved ability to predict is sometimes transposed into a belief that they can actually "influence" or control the outcome of a gamble. These considerations shift the focus of attention towards another notion which appears crucial in hazard assessment and which has been touched upon earlier; assessment of outcomes and consequences.

3.6 ASSESSMENT OF OUTCOMES

According to Ross (1974), assessment of potential outcomes is just as important as assessment of likelihood in hazard appraisal. She writes:

> "Even if agreement were reached on the probability of an accident, people might still disagree about the danger. Some might feel that the accident is trivial, others that it is dreadful. Some regard death as the ultimate disaster, while others regard it as the gateway to a better life. Some feel that maiming is worse than death, others that life of any sort is previous" (Ross, 1974).

3.6.1 Punishment of crime

Cohen (1970a) cites Lord Gardiner as having once said that even if all punishment were abolished it would have no effect on the rate of crime. What counted, Lord Gardiner remarked, was the likelihood of getting caught (or of getting away with the crime).

Cohen himself refutes the judge's argument on the grounds that likelihood of getting caught combines with a person's perception of the punishment to form the perceived value of risk when he considers committing a crime. Cohen states that two considerations must be pointed out in relation to outcome severity assessment.

Firstly, the graph which best describes the effect of increased severity on deterrence is an inverted U-shape curve. In other words, up to a certain point, the greater the severity, the greater the deterrent effect of punishment. After this point severity reaches maximum efficiency; beyond this point increasing severity gradually decreases the deterrent effect.

The second consideration mentionned by Cohen is that some criminals do assess the risk associated with their crime before committing it. However others take neither likelihood of getting caught nor severity of potential punishment into consideration before committing crime. Such is the case, for instance, of those shop-lifters who act on the spur of the moment.

There are qualitative differences between committing crime and undertaking a physically dangerous activity. But a limited analogy may be drawn to suggest that severity of pun'shment is to crime what dreadfulness of outcome is to a hazard.

3.6.2 Consequences of accidents

There are indications that considerations similar to those mentioned by Cohen (1970a) can be raised in relation to hazard assessment. For instance dangerousness does have a deterrent effect. This is exemplified in the results of research by McKenna (1978) where a majority of respondents stated their unwillingness to have a job more dangerous than their present one. Like for punishment, however, this effect can actually decrease. For example, Kates (1967) mentions areas of relatively high risk of violent storms;

most inhabitants of such areas would rather face the consequences of storms than move to areas of lower risk. Their unwillingness to move could also indicate that they feel confident in their ability to cope with the consequences of storms.

There are also cases where dangers are not considered at all. For example, Hale and Pérusse (1978) mention workers who said that if they stopped and considered all the hazards around them they would be too worried to carry out any work.

3.6.3 Subjective value

Dreadfulness, like severity of punishment, may not have the same meaning or the same value for everyone. This argument is put forward in the quotation from Ross (1974) in section 3.6 of this chapter. It is also possible that some people tend to underestimate the severity of outcomes and that such people may be "risk-takers". For instance Robaye et al (1963) mention that accident repeaters tend to adopt riskier behaviour and to underestimate seriousness of outcome. Which (of accident experience or severity underestimation) is cause and which is effect is unclear.

A number of variables are reputed to influence perceived severity of accident outcomes. Glendon (1976) mentions that the more people are killed in an accident the more dreadful the accident is perceived to be. He also mentions that if victims are defenseless people (e.g. children, old-age pensioners) perceived dreadfulness is higher. Certain types of deaths (e.g. slow painful deaths) and certain fates (e.g. "paralysis from the neck down"; Green and Brown, 1976 a) tend to be perceived as particularly dreadful.

3.6.4 Effects of experience

It was pointed out earlier that accident experience influences peoples' assessment of the likelihood of accidents. There is also evidence which suggests that suffering from the consequences of certain dangers alters the perceived dreadfulness of those dangers.

Green and Brown (1976a) observe that some respondents in their research, having sustained certain injuries (e.g. broken limbs), tended to rate these injuries as less dreadful than did respondents who had not sustained these same injuries. On the other hand, some respondents rated other injuries they had sustained (e.g. eye injuries) as more dreadful than did those who had not sustained the same injuries.

Therefore there are grounds for believing that injury experience does have an influence on assessment of dreadfulness. At present, however, there is insufficient research information for the exact nature of that influence to be fully understood.

3.6.5 Control, likelihood and dreadfulness

Ross (1974) argues that, even when likelihood and dreadfulness are perceived in the same way by everyone, "further disagreement can arise when people try to compare a highly probable but small disaster with an improbable but serious one". In this case both assessment of likelihood and assessment of dreadfulness would combine and interact to form a composite assessment of a hazard.

As pointed out earlier such interaction appears to underlie some results in the studies by Melinek et al (1973) and by McKenna (1978). It was also pointed out that what appears to be the main link between the two is the concept of control, be it control over occurrence or control over consequences. Both types of control appear to be associated in people's minds.

Although some hypotheses have been put forward to describe the nature of perceived likelihood and perceived dreadfulness, little is known about these two concepts. But it appears that ev less is known about the concept of control. What empirical evidence there is about assessment of probabilities, assessment of dreadfulness and perceived control is reviewed in the next chapter.

3.7 SUMMARY

It is argued at the beginning of this chapter that hazard assessment is a form of recognition of danger. But it is also argued that, although it does involve some decision-making, it is different from other types of decision-making, for instance in terms of risk-taking.

Despite important differences between gambling and facing danger, it may be seen that certain mechanisms underlying gambling can also be identified in hazard assessment. For example, although assessment of probabilities is important, it is not the only component of hazard assessment. Other variables, such as accident experience and belief in one's own skill and knowledge, have a bearing upon perceived likelihood of occurrence.

Another important aspect of hazard assessment, it is argued, is assessment of dreadfulness. Some of the same variables which influence perceived likelihood, it is suggested, also influence perceived dreadfulness.

One variable which appears to form a key link between perceived likelihood and perceived dreadfulness is perceived degree of control over events. But relatively little has been written on these three aspects of hazard assessment. Empirical observations must be turned to for further indications about the nature of hazard assessment.

CHAPTER 4: REVIEW OF RESEARCH FINDINGS

Cohen (1970 a) and Ross (1974) suggest that there are three major dimensions of hazard assessment. Both authors mention that perceived likelihood of occurrence of an unwanted outcome is one dimension, and they agree that perceived dreadfulness of outcomes and perceived control over events are the other two.

Their arguments, however, are mostly speculative. Supportive evidence is either anecdotal or derived from studies in other areas (e.g. risk-taking and crime punishment). Studies directly relevant to hazard assessment have been carried out by Golant and Burton (1969), Green and Brown (1976 a, 1976 b, 1977 a, 1977 b), Champion (1977), and by Fischhoff et al (1978). This chapter is a review of the relevant findings in these studies.

4.1 THE USE OF THE SEMANTIC DIFFERENTIAL TEST

Golant and Burton (1969) used the semantic differential test in asking respondents to evaluate 12 hazards (listed in Table 4.1). The researchers also devised 21 seven-point scales (or bi-polar concepts) for rating each hazard. These scales are listed in Table 4.2 .

TABLE 4.1

LIST OF HAZARDS IN GOLANT AND BURTON'S SEMANTIC DIFFERENTIAL TEST

1. Air pollution
2. Auto accident
3. Boat accident
4. Building collapse
5. Earthquake
6. Epidemic
7. Flood
8. Housefire
9. Riot
10. Snowstorm
11. Tornado
12. Water pollution

TABLE 4.2

LIST OF SCALES DEVISED BY GOLANT AND BURTON (1969) TO EVALUATE HAZARDS IN THEIR SEMANTIC DIFFERENTIAL TEST

1.	Passive	Active
2.	Orderly	Chaotic
3.	Natural	Unnatural
4.	Stable	Unstable
5.	Widespread	Localized
6.	Peaceful	Ferocious
7.	Fair	Unfair
8.	Dissonant	Harmonious
9.	Slow	Fast
10.	Strong	Weak
11.	Private	Public
12.	Important	Unimportant
13.	Relaxed	Tense
14.	Erratic	Periodic
15.	Determinate	Fortuitous
16.	Yielding	Tenacious
17.	Artificial	Natural
18.	Controllable	Uncontrollable
19.	Pleasant	Unpleasant
20.	Light	Heavy
21.	Constrained	Free

The test scales were assembled randomly both as a precaution against response bias and to avoid monotony. The questionnaire was administered "to 58 subjects, primarily University summer extension students of various socio-economic backgrounds" (Golant and Burton, 1969). Matrices of correlation coefficients were calculated both between scales and between hazards from means of the 58 individual ratings. The matrices were then processed by principal component analysis.

4.1.1 Three types of hazards

Three components were extracted from the hazards matrix. Golant and Burton (1969) call the first component "man-made hazards". It explains 34.3% of the test variance and six items load significantly on it: housefire, building collapse, boat accident, auto accident, riot and epidemic.

The second component accounts for 24.4% of the test variance. Four items load significantly on it: flood, tornado, earthquake and snowstorm. Golant and Burton (1969) label this component "natural hazards".

Only two items load significantly on the third component: air pollution and water pollution. In view of these, Golant and Burton (1969) call the component "quasi-natural hazards". It explains 18% of the test variance.

4.1.2 Four dimensions of hazard assessment

The components which were extracted from the hazards matrix might deserve further consideration. For instance the labels which Goland and Burton (1969) choose for them could be interesting topics of discussion. But other findings from the same research are more relevant to the main theme of this thesis. These findings are related to the four components which were extracted from the scales matrix. Between them the components explained 45.8% of the test variance (18%, 12.4%, 8.9% and 6.5% respectively).

Golant and Burton (1969) affix the label "stability" to the first component. The following scales loaded significantly on it: passive-active, orderly-chaotic, stable-unstable, peaceful-ferocious, dissonant-harmonious, slow-fast, relaxed-tense, and pleasant-unpleasant.

This component was labelled "stability because it contains a predominance of adjectival scales depicting various states of equilibrium or deviations from some *normal* condition" (Golant and Burton, 1969). Although this argument has face validity, a closer examination of the results reveals inconsistencies both in the polarity of some scales and in the signs of their loadings. Further ambiguity about the label arises from the fact that the scales which purports to measure the concept of "stability" has the lowest significant loading of eight scales.

"Controllability" is the label chosen by Golant and Burton (1969) to describe the second most important factor. It comprises the scales: natural-unnatural, fair-unfair, artificial-natural and controllable-uncontrollable. As for the previous component there are indications that "Controllability" may not be the most appropriate description for this component. For instance, the scale controllable-uncontrollable has only the third highest loading out of four scales. Furthermore the two scales which account for the highest loadings both have the adjective "natural" as one of their poles.

It may be that a label such as "man-made hazard/act of God" would be more befitting for this component. Such a dimension appears to predominate the factors extracted from the hazards matrix. In addition it is possible that controllability is salient only inasmuch as natural disasters have the connotation of being uncontrollable and man-made hazards are generally deemed to be controllable. Similar experiments using the relevant scales could shed some light on this point.

Golant and Burton (1969) call their third component "magnitude". The scales which load significantly on it are widespread-localized, strong-weak, private-public, important-unimportant, determinate-fortuitous, yielding-tenacious and light-heavy. This component might be associated with the concept of dreadfulness. Unfortunately no scale in Golant and Burton's test conveyed a direct meaning of dread (e.g. fatalities, injuries, damage, etc.). A link between magnitude and dreadfulness could be investigated only with the use of a test incorporating scales measuring both concepts.

The fourth component extracted from the scales matrix is labelled "expectancy" by Golant and Burton (1969). Comprising it are the scales: erratic-periodic and free-constrained. It is possible that this component refers to perceived probabilities of occurrence. If that is the case the salient concepts appear to be absent from the test.

In the light of earlier discussion, it might be expected that the concept of perceived likelihood would emerge as relatively important. This fourth component, however, was the last significant one to be extracted. It must be pointed out that none of the scales devised by Golant and Burton dealt with the concept of likelihood. It may be that the presence of such scales could have made this component more salient. It it is assumed that the relevant scales are absent from the last three components, then Golant and Burton's results appear to substantiate the views of Cohen (1970 a) and Ross (1974) to the effect that perceived control over events, perceived dreadfulness and perceived likelihood of occurrence are major dimensions of hazard assessment.

An additional dimension emerged from Golant and Burton's results which was not discussed earlier. The "stability" component explains more test variance than any other. But at the same time, it is the most ambiguous one and the one whose label is most dubious.

4.1.3 Methodological limitations

A number of methodological considerations preclude further interpretation of Golant and Burton's results. For instance these authors tested only one population which they treat as homogeneous. They are aware of this limitation, and suggest that replication of the study could prove a very profitable avenue of research for exploration.

Other weaknesses of their method are related to statistical analysis. The researchers used only principal component analysis to extract underlying trends or dimensions. Child (1976) suggests that the use of only one factor programme makes testing the "robustness" of factors impossible. Child (1976) further suggests that provisions should be made to take into account unique variance in factors which explain little test variance. As the four factors extracted from the matrix of scales between them explain less than half the test variance, the extent to which they are "polluted" is not known.

Ambiguity in components may also be attributable to ambiguity in the rating scales. There are a number of indications that only parts of rating scales were used by respondents:

1) mean ratings for hazards "varied only from 4.02 to 4.40" (Golant and Burton, 1969);
2) for 9 of the 21 scales only 2 of the 5 points were used;
3) for another 8 scales only 3 of the 5 points were used;
4) on 13 of the 21 scales two thirds (8) or more of the hazards were rated on the same point.

Although the statistical significance of these indications cannot be assessed without the raw data, they seem to point towards a limited use of some scales. Such small variance usually means low discriminability (Guilford and Fruchter, 1973), which in turn may lead to unbalance in the relative importance of factors (Slater, 1972).

The scales chosen by Golant and Burton may have caused another problem which is not given due consideration by the researchers. Some of the scales may have been irrelevant in assessing certain hazards. Osgood et al (1957) point out that irrelevant scales can be rated at mid-point values (e.g. a 4 on a seven-point scale). Bannister and Mair (1968) however point out that it is erroneous to use the same point to represent both an intermediate rating and irrelevance. They also mention that this can lead to serious bias in the results.

The scales which respondents used to rate hazards were devised by the researchers. It is therefore possible that, as postulated earlier, some salient concepts were overlooked. If that is the case further experiments which would incorporate such possible concepts might show important modifications in the pattern and relative weight of components.

It can be argued that these and similar problems and constraints are associated with most rating scales. If that is the case interpretation of results must take such constraints into account. But two studies, described later, have used repertory grids in preference to semantic differential tests, and problems usually associated with the latter appear to have been avoided.

All these methodological considerations tend to restrict further interpretation of Golant and Burton's results. Therefore, although they appear to lend support to other authors' views on hazard assessment, these results on their own cannot be considered to be fully reliable.

4.2 THE MEASUREMENT OF PERCEIVED SAFETY

Research by Green and Brown (1976 a, 1976 b, 1977 a, 1977 b) provides some answers to the questions raised above. Green (1975 b) describes the aims of their research as being:

1) to derive a measure of personal safety;
2) to define an acceptable level of safety; and
3) to define aversion as a measure of safety.

To attain goals 2 and 3, a measure of safety had to be devised. In other words, goal 1 was a prerequisite to the other two goals. Thus, as Green (1975 b) points out, these goals are sequential.

Four steps, in the form of experiments (named E1 to E4) were designed to reach the three goals. The purpose of E1 was to find out if dimensions used by individuals in the assessment of hazards could be identified. For these dimensions to be validated and used in subsequent experiments it had to be shown that they were capable of a certain level of discrimination. In E2 factors identified in E1 were used to "calibrate the hazards" and to identify respondents' perception of the safety of the various situations and activities identified in E1 as hazards.

E3 was devised to test the stability of the measure of safety elaborated in E2. The measure had to show test-retest consistency. The experiment also sought whether the measure fluctuated coherently as a function of additional information (correct in one experiment and incorrect in another) presented to respondents about the hazards. The purpose of E4 was to "determine the utility to the individuals of their personal safety" (Green and Brown, 1977 c) and to assess individual satisfaction with the safety levels of those activities and situations whic were used and compared in the other experiments.

This series of experiments has provided a wealth of interesting results. However, those results which are relevant to the argument within this thesis are the main findings of E1 and E2. Therefore the methodology and results of these two experiments will be reviewed in greater detail. Further information about E3 and E4 is readily available in the results summaries provided by Green and Brown (1977 a, 1977 b, 1977 c) at the beginning of their reports on these experiments.

4.2.1 The use of the repertory grid technique

To discover the criteria used by people in assessing hazards, Green (1975 b) advocates the repertory grid technique. This technique is discussed in greater detail in Chapter 5. Unlike the semantic differential test (e.g. in Golant and Burton, 1969) the repertory grid enables participants themselves to select the criteria (or constructs, as they are hereafter called). This has the advantage that respondents use rating scales which have some psychological meaning for them.

Because it "introduces interdependencies between the concepts being examined and the scales" (Green, 1975 b) on which they are rated, use of the repertory grid ensures that constructs are relevant to the concepts being rated. As scales in a semantic differential test are chosen by the researcher , their relevance is difficult to assess Some of the problems incurred through lack of relevance are discussed in the review of Golant and Burton's results.

"A subsidiary advantage of the repertory grid is that it may throw up some further attitudinal dimensions" (Green, 1975 b) which might be important but forgotten or overlooked by the research (Bannister and Mair, 1968). As argued earlier such oversights may have caused shortcomings in Golant and Burton's research.

Table 4.3 lists the 21 hazards (hereafter refered to as elements) which were used as stimuli in El. Table 4.4 contains

TABLE 4.3

LIST OF ELEMENTS USED BY GREEN AND BROWN (1976 a) IN EXPERIMENT E1

1. Hotel fire
2. Coal mining accident
3. Air pollution
4. Fire in a discotheque
5. Illness
6. Train crash
7. Accidental release of nuclear radiation
8. Home fire
9. Plane crash
10. * Earthquake
11. * Car crash
12. Rock-climbing accident
13. * Accident on a building site
14. Being struck by lightning
15. Accident in a chemical plant
16. Skiing accident
17. Factory fire
18. Accident in the home
19. Food poisoning
20. Motorcycle crash
21. Being knocked down while crossing the road

* Indicates an element similar or identical to one used by Golant and Burton (1969).

TABLE 4.4

LIST OF CONSTRUCTS MOST OFTEN ELICITED IN E1 AND THE NUMBER OF RESPONDENTS WHO MENTIONED THEM

	Poles of the construct		Frequency
1.	Self control	- Out of own control	16
2.	Act of man	- Act of nature/God	13
3.	Necessary	- Unnecessary activity	13
4.	Controllable	- Uncontrollable	12
5.	Avoidable	- Unavoidable	12
6.	Preventable	- Unpreventable	12
7.	Frequent	- Infrequent	12
8.	Easy to escape	- Difficult to escape	10
9.	Minor	- Lethal	10
10.	Rely on other people	- Rely on self	8
11.	Scaring	- Not scaring	7
12.	High risk of accident	- Low risk of accident	7
13.	Large consequences	- Small consequences	7
14.	Most dangerous	- Least dangerous	6
15.	Aware of danger	- Unaware of danger	6
16.	Many killed	- Few killed	6
17.	Blame assignable	- No blame assignable	5
*18.	Slow event	- Fast event	5
19.	High risk of death	- Low risk of death	5
20.	Traditional danger	- Modern danger	4

* Indicates a construct similar or identical to a scale used by Golant and Burton (1969).

a list of the constructs which were elicited most often, and their frequencies, in response to the elements in Table 4.3.

Seventeen students from the School of Architecture in the University of Dundee participated in E1. Rating of elements on constructs was performed by all respondents.

The aim of E1 was to provide the list of criteria which people used in assessing hazards. A study of the underlying structures and of the inter-relationships between constructs was not the main purpose of the experiment. Therefore individual grids were only scored and no detailed analysis of the repertory grids was performed.

In E2 Green and Brown (1976 b, 1977 b) measured the perceived dangerousness and the perceived safety of the various elements. They also devised questions asking respondents to assess the elements on some of the constructs listed in Table 4.4 as well as on different constructs which other researchers had suggested were important in the assessment of levels of safety. Correlations were calculated between perceived dangerousness perceived safety and all the constructs used in E2 in order to find out which constructs were the best predictors of the two types of perceptions.

A few elements (e.g. illness and electric shock) had to be rejected after pretesting "as the respondents stated that their inclusion with situational hazards was an apples and oranges one" (Green and Brown, 1976 b). The main problem with these hazards was that the researchers "wished... to define each hazard in situational terms whether there was both the possibility of that hazard occurring and, should it occur, the chance of being harmed by it" (Green and Brown, 1976 b); but some hazards lent themselves to formulations which could include either only the possibility of occurrence or only the chance of being harmed.

4.2.2 Discussion of the results of E1 and E2

Green and Brown (1977 b) noticed that respondents appeared to categorize bodily harm into two classes: 1) minor injuries and 2) death and death-like (i.e. nearly-as-bad-as-death and "supra-lethal" injuries) consequences. Respondents rated "bruises" and "sprained ankle" as being hardly of more consequence than being unhurt by an accident. In contrast they rated two fates ("paralysis from the neck down" and "brain damage") as being worse than death; other types of injuries (e.g. "loss of sight of both eyes") were rated as being nearly as severe as death.

There are indications that, when participants in E2 assessed likelihood of injury, in fact they assessed the likelihood of trivial consequences. When likelihood of death was assessed, the notion of "death" appeared to encompass very severe and supra-lethal injuries as well. Furthermore, "the evidence suggests that respondents self-define an accident as a near or supra-lethal accident and do not include less serious injury" (Green and Brown, 1977 b).

Another important finding from E2 is that "perceived safety and perceived hazardousness are not identical concepts" (Green and Brown, 1977 b). In other words the researchers found that perceived safety may not be the opposite of perceived dangerousness. The main characteristics of perceived safety appear to be "perceived probability of accident occurrence" and "perceived involuntariness of the activity". Whereas "perceived probability of accident occurrence" is also a characteristic of perceived dangerousness, the other main trait of this dimension is "perceived degree of personal control over risk assumed".

The variable P $(D|A)$, the conditional probability of death should an accident occur, was correlated significantly ($p < .01$) with perceived safety. However, since it appears that respondents considered as being accidents only those events which had very serious consequences, assessing the likelihood of an accident almost automatically implied assessing the likelihood of death. Thus P $(D|A)$ was not a predictor of perceived safety.

Perceived hazardousness appears to refer both to accident occurrences and to control over them. Some constructs which conveyed both notions were correlated with both factors. Thus it seems that likelihood is perceived as being a function of perceived control rather than as a function of specific indicators like P $(D|A)$. This tends to reinforce an earlier argument that subjective probabilities are more important than objective probabilities in hazard assessment.

As pointed out earlier, one of the main predictors of perceived safety in E2 was "perceived involuntariness of the activity". Using the method of "revealed preferences" Starr (1969) postulates that for equivalent levels of benefits people are prepared to accept higher levels of risk for activities voluntarily engaged in than for involuntary activities. Starr's analysis has been much criticised (Green, 1974; Fischhoff et al, 1978). Nevertheless Fischhoff et al (1978), using a different approach, found that for equivalent levels of benefits levels of acceptable risk were higher for voluntary activities.

Why perceived voluntariness is a good predictor of perceived safety is not clear. One possible explanation is based on the concept of control. It may be that activities voluntarily engaged in are perceived as being under the actor's control. Assuming that a person thinks that he is in control of events, confidence in his own ability to cope may then, as argued in the previous chapter, become a source of hazard underestimation.

Such an explanation of the role of perceived voluntariness in hazard assessment is in fact a series of hypotheses based on the concept of "perceived degree of personal control over risk". However in E2 perceived control was a predictor of perceived hazardousness only and not of perceived safety.

Other findings of E2 suggest an alternative explanation. Green and Brown (1977 b) mention that perceived control and perceived voluntariness were significantly ($p < .001$) correlated.

Furthermore it appears that perceived hazardousness is assessed in terms of likelihood of a mishap whereas perceived safety is assessed in terms of the severity of potential accidents. Thus it could be that perceived voluntariness reflects perceived control over consequences whereas "perceived degree of personal control over risk" reflects perceived control over events potentially leading to an accident. The fact that the construct "easy/difficult to escape" was mentionned 10 times might be an indication that control after the event is an important consideration in respondents' minds.

Thus the results of E1 and E2 suggest that hazards are assessed in two stages. The first stage is to assess "hazardousness", i.e. likelihood of an accident. During the second stage "safety" is assessed; in other words the severity of potential consequences (and perceived control over consequences) is taken into consideration in the second stage of hazard assessment.

Some results of E1 appear to support this hypothesis. For instance the notions of "controlability", "preventability" and others related to limitation of undesired outcomes were significantly correlated with "perceived hazardousness"; these notions were reflected in the most frequently mentioned constructs in E1. It can be argued that the constructs were frequently elicited mainly because the elements used strongly suggested the corresponding notions to respondents. However, it can also be argued that respondents have mentioned quite a few other characteristics of hazards; if the predictors of perceived hazardousness were mentioned more often this could suggest that these characteristics are prominent criteria in assessing hazards.

Notions related to dreadfulness (e.g. scariness, severity of possible injury and size of consequences) are elicited less frequently. The possibility that elements did not suggest the notion of dread can be ruled out since quite a few respondents did mention constructs related to dreadfulness. It may be that the notion of dreadfulness is simply less important than the notion of control. This, however, appears to contradict Cohen's (1970 a) and Ross's (1974) arguments.

Two explanations can be suggested which could account for the differences in construct frequencies. The first of these possible explanations relates to an argument discussed by Hale and Pérusse (1978). They postulate that thinking about harm provokes stress and that, as a result, some cognitive dissonance mechanism might obliterate consideration of dreadful consequences. The second explanation refers to the two-stage assessment mechanism discussed earlier. It could be that dreadfulness of potential consequences is assessed only for those hazards which are deemed to be uncontrollable. This would imply that assessment of dreadfulness is carried out <u>after</u> assessment of degree of control. Thus, if both assessments are carried out in sequence rather than in parallel, the first mechanism in the sequence may well appear to be prominent over the other.

4.2.3 Limitations upon interpretation

A number of considerations limit the conclusions which can be drawn from experiments E1 and E2. Firstly, as the experiments were progressing, a few elements were either added to or removed from the original list. Secondly, the wording of some elements was altered at various stages of testing. It is therefore not known if the results may be generalized to all the elements listed in Table 4.3. Thirdly, the purpose of the experiments was to devise a measure of perceived safety; therefore the experiments were not designed to measure the relative importance of the various factors which underlie assessment of hazards. Fourthly, some elements depicted relatively rare occurrences; for example "only one person (has) died of snakebite in Britain in the past 50 years" (Green and Brown, 1976 b). In contrast elements like "house fire" depicted relatively more frequent occurrences. Moreover some elements (e.g. "earthquake") depicted more specific events than elements like "accident in the home". The latter is a general category which covers a whole array of different mishaps. Similar differences also exist in the elements used in Golant and Burton's (1969) study.

Given such disparities between elements it is very likely that certain constructs could not be used to rate some elements. It may also be that, because of "interdependencies between the concepts...and the scales" (Green, 1975 b), some constructs which might have emerged, if the elements had all been similar types of situations, were precluded by the diversity of elements.

Fifthly, again because of the specific aims of the experiments, some constructs elicited in E1 (e.g. whether blame could be apportioned) were irrelevant in E2. A large proportion of the factors which were retained for E2 were possible indicators both of perceived likelihood and of perceived dreadfulness. Such factors have led to interesting findings discussed in this section.

Only those results of E1 and E2 which are relevant to the argument of this thesis were looked into. Their interpretation remains largely hypothetical. Green and Brown (1976 b, 1977 b, 1977 d) point out on a number of occasions the need to know more about the basic mechanisms of hazard assessment.

4.3 REPERTORY GRIDS ON INDUSTRIAL HAZARDS

The main purpose of Champion's (1977) research was to achieve a better understanding of the way in which industrial hazards are assessed. Champion remarked that the elements used by Golant and Burton (1969) and by Green and Brown (1976 b) were fairly heterogeneous. He made the assumption that situation-related and more relevant criteria would be elicited if the hazards (elements) were more homogeneous. He therefore used only industrial hazards as elements.

For much the same reasons as Green (1975 b), Champion (1977) used the repertory grid as his main measurement instrument. In both Golant and Burton's (1969) study and Green and Brown's (1976 b) research, elements to be rated had been provided for the respondents by the researchers. But one variant of the repertory grid technique suggests that elements should be chosen within the respondent's range of familiarity (Kelly, 1955). Therefore Champion (1977) chose hazards from the respondents' usual work environment. The opportunity of doing so was presented to him as part of a series of experiments in which I was involved (Pérusse, 1978).

4.3.1 Element elicitation: the diary technique

The aim of the experiments was to assess the reliability of the diary technique as an instrument to facilitate hazard spotting and reporting. The diary technique usually requires participants to keep a record, or diary, of their behaviours or observations of a specific type over a period of time. For these experiments record cards, called Hazard and Accident cards (HAC) were devised.

Participants were asked to carry with them one card each day for 5 weeks. They were also asked to enter on the HAC details of hazards and accidents which they came across during the course of their normal activities. Hazards and accidents were to be entered no matter where they occurred - at home, on the road, at work or elsewhere. Participants were instructed not to go out of their way to spot hazards and accidents, but were told to enter those they saw as soon as possible after noticing them. All participants were issued with a set of written instructions.

One of the experiments was run in a beer kegging plant near Runcorn, Cheshire. This experiment was reaching completion when Champion (1977) undertook his research. It was therefore decided to use the entries from that exercise which were related to industrial hazards. The 19 participants who completed the exercise provided 19 such hazards. These were put together to form the list of elements for the repertory grid. These elements can be found in Table 4.5.

TABLE 4.5

LIST OF ELEMENTS USED IN CHAMPION'S (1977) REPERTORY GRIDS

1. Main valves left open on oxy/acetylene welder
2. Donkey jacket draped across convector heater
3. Emergency doors blocked-out for film show
4. Pipes across walkway
5. Badly fitting grid sticking up above floor level
6. Electric cable lying on floor
7. Tablet of soap on wet floor
8. Gas cylinder left free-standing on ramp
9. Scaffolding stacked in dangerous manner
10. Forklift truck being driven without warning light
11. Guard missing off machinery
12. Beer on floor makes it slippery
13. Operator leaning across conveyor to operate mach   ry
14. Steam gushing out of tanker where people walk
15. Cut-off electric eye not working
16. Oil spillage
17. Kegs thrown off conveyor by faulty reject arm
18. Keg lift faulty so operator has to brake hard
19. Bad ladder footing

4.3.2 Construct elicitation

Only one further visit to the kegging plant was possible because of the summer holiday period. Therefore the visit could be used to get the respondents either to elicit the constructs or to score the grids, but not to do both. Champion (1977) decided that the final visit to the plant would be used for grid scoring. Therefore he undertook to conduct a series of 5 pilot experiments, using a range of different people. During these experiments constructs were elicited in group discussions by a total of 14 respondents of both sexes and of various background and occupations. The stimuli, or elements, presented to the respondents for construct elicitation were those to be used for the grids in the main experiment.

Those dimensions which were mentioned in at least two of the pilot studies were selected to form the list of constructs used in the grids. Champion (1977) opted for a five-point scale anchored at both ends. In other words the two extremes of each construct were supplied, corresponding to scores of 1 and 5 respectively. Table 4.6 presents the final list of the 17 constructs which were retained. The poles on the left-hand side of Table 4.6 correspond to the end of the scale scored as 1.

4.3.3 Grid scoring and analysis

Elements and constructs were arranged in questionnaire form. Each page of the questionnaire contained the list of elements and a five-point scale for each of them, and was headed by one construct. To avoid response pattern biases, the order of the pages

TABLE 4.6

LIST OF CONSTRUCTS USED IN CHAMPION'S (1977) REPERTORY GRID

1.	Very often encountered in my job	- Never encountered in my job
2.	Necessary result of the process	- Danger not a necessary result of the process
3.	Temporary danger	- Permanent danger
4.	Moving danger	- Stationary danger
5.	Easily spotted danger	- Very difficult to identify danger
6.	Danger is immediately present	- Danger is dependent upon other things
7.	Very likely to cause an accident	- Very unlikely to cause an accident
8.	Easy to avoid consequences of danger	- Impossible to avoid consequences of danger
9.	Only a trivial injury	- Permanent disability
10.	Very likely to kill	- Very unlikely to kill
11.	Only one person at risk	- Every person in the plant at risk
12.	Preventable	- Unpreventable
13.	Takes a specialist to put it right	- Anyone can put it right
14.	Danger arises from bad design feature	- Danger has nothing to do with design
15.	Management's fault	- Nothing to do with management
16.	Operator's fault	- Nothing to do with operator
17.	Due to inadequate training	- Danger has nothing to do with training

(i.e. of the constructs) was different for each respondent. Twelve respondents from the kegging plant scored the grids. Each respondent was briefed individually. Scoring took an hour on average.

Each grid was processed by a computer program called Ingrid 72 and devised by Slater (1972). The main features of the program are described in Chapter 6.

Twelve participants each rating 19 elements on 17 constructs gives a maximum of 3876 ratings. But 5 of these ratings were missing. As Ingrid 72 does not take missing values into account these were replaced by the population's mean score for those cells. Then a consensus grid was computed. Each cell of the consensus grid was the arithmetic mean of the corresponding cell in the 12 individual grids. The consensus grid was subsequently processed in the same way as the individual grids.

Champion (1977) focused his attention on factor analysis (principal components solution, no rotation of factors) of the consensus grid. The Bartlett test which was performed extracted 4 significant factors. Champion chose a criterion loading of $\pm$ 0.3 to decide which constructs loaded significantly on those factors. The resulting components are summarized in Table 4.7, in which the significant items (constructs) and their loadings are listed for each component.

TABLE 4.7

SUMMARY OF THE PRINCIPAL COMPONENTS EXTRACTED FROM CHAMPION'S (1977) CONSENSUS GRID

Component 1		Component 2		Component 3		Component 4	
"Scope for personal control"		"Injury Potential"		"Familiarity"		"Danger circumstances"	
Item	Loading	Item	Loading	Item	Loading	Item	Loading
8	-0.91	4	-0.74	9	+0.63	7	-0.81
14	+0.90	10	-0.74	5	-0.60	11	-0.65
16	-0.90	9	+0.72	1	-0.58	15	+0.59
2	+0.79	13	+0.66	3	+0.56	6	-0.40
17	-0.75	1	+0.65	6	-0.47		
12	-0.69	11	-0.52	17	-0.47		
15	+0.68	5	+0.45	16	-0.38		
13	+0.63	12	-0.39				
3	-0.58						
6	+0.44						
10	-0.39						
4	+0.37						
% variance: 34.6		% variance: 20.2		% variance: 13.4		% variance: 12.2	

4.3.4 Discussion of results

It is interesting to note that the factor in Champion's results which explains the largest amount of variance seems to convey the notion of perceived degree of control over the situation. As for the second most important factor, it appears to convey the notion of dreadfulness.

The third and fourth factors are more ambiguous. It is possible, for instance, that "obviousness of danger" could be as good a label for the third factor as "familiarity"; this factor has overtones of salience, of importance of hazards, of something to be taken notice of. Finally ambiguity in the label "danger circumstances" probably reflects ambiguity in the fourth factor itself. It could be that labels such as "environmental riskiness" or "person's fault" could also be affixed to it.

When the results of principal component analysis are printed out by Ingrid 72, element as well as construct loadings are listed for each factor. For each component or factor in Champion's study an analysis of the nature of those elements whose loading was significant would have yielded further indications about the nature of the component. A comparison of positively and negatively laden elements would have further clarified the meaning of each pole of a component. Unfortunately, Champion did not analyse element loadings.

4.3.5 Some limitations

There are methodological considerations which make it difficult to accept Champion's factors as they stand. Firstly, as pointed out earlier in relation to Golant and Burton's (1969) study, in component methods unique variance becomes merged with common variance to an increasing extent as more factors are extracted. In Champion's analysis no allowance was made for unique variance. However Champion's components explained more common variance than did Golant and Burton's; therefore, it is likely that unique variance intruded less in Champion's results than in Golant and Burton's. Secondly, Champion only relied upon principal component analysis to extract underlying trends. As there is no validation of those factors by other methods, it is not known how "robust" they are. Thirdly, each cell of the consensus grid was the arithmetic mean of the corresponding cells in the 12 individual grids. But arithmetic means may not always be the most accurate representation of central tendencies (Guilford and Fruchter, 1973). There may have been inadequacies in the matrix of arithmetic means which the researcher did not control for.

4.4 FURTHER EVIDENCE ON TWO DIMENSIONS

The main purpose of Fischhoff et al's (1978) research was to investigate the mechanism for deciding on the acceptability of safety levels. To do so, they asked the participants in their study to evaluate "each of 30 different activities and technologies with regard to (1) its perceived benefit to society; (2) its perceived risk; (3) the acceptability of its current level of risk; and (4) its position on each of nine dimensions of risk" (Fischhoff et al, 1978). The activities and technologies thus evaluated are listed in Table 4.8.

The rating of the various activities and technologies on the "nine dimensions of risk" was done on seven-point scales. Thus this particular task was akin to rating a semantic differential

TABLE 4.8

LIST OF ACTIVITIES AND TECHNOLOGIES EVALUATED BY RESPONDENTS

IN FISCHHOFF ET AL'S (1978) RESEARCH

1) Alcoholic beverages
2) Bicycles
3) Commercial aviation
4) Contraceptives
5) Electric power (non nuclear)
6) Fire fighting
7) Food coloring
8) Food preservatives
9) General (private) aviation
10) Handguns
11) High school and college football
12) Home appliances
13) Hunting
14) Large construction (dams, brides, etc.)
15) Motorcycles
16) Motor vehicles
17) Mountain climbing
18) Nuclear power
19) Pesticides
20) Power mowers
21) Police work
22) Prescription antibiotics
23) Railroads
24) Skiing
25) Smoking
26) Spray cans
27) Surgery
28) Swimming
29) Vaccinations
30) X-rays

test. The dimensions (or scales) used are listed in Table 4.9.

4.4.1 Some results

There were two groups of respondents, each performing a different combination of some of the four tasks described earlier. Both groups, however, rated the activities and technologies on the nine scales. It was found that both series of results were similar enough to be pooled. Furthermore there was good interparticipant agreement about the various ratings.

The researchers wanted to identify the "basic dimensions of risk underlying the nine characteristics". Therefore they performed a principal component analysis. Again the results from the two sets of data "were so similar...that they were averaged". A varimax rotation was performed on the resulting factors, but the researchers mention that "it produced no improvement in interpretability". Table 4.10 lists the loadings of the nine characteristics on the two extracted factors.

The most important factor correlated significantly with all risk characteristics except severity (i.e. certainty of death should an accident occur). The authors chose the label "Technological risk" for this factor. The second factor correlated significantly with three characteristics: severity (certainty of death), common-dread (i.e. whether people have learned to live with a certain risk) and chronic-catastrophic (i.e. the number of people which a single event can kill). Therefore the label "severity" was chosen for this factor.

TABLE 4.9

LIST OF DIMENSIONS ON WHICH PARTICIPANTS IN THE RESEARCH BY FISCHHOFF ET AL (1978) RATED ACTIVITIES AND TECHNOLOGIES

1)	Voluntary	-	involuntary
2)	Immediate effect	-	delayed effect
3)	Known precisely by person exposed	-	not known
4)	Known precisely to science	-	not known
5)	Uncontrollable	-	controllable
6)	New	-	old
7)	Chronic	-	catastrophic
8)	Common	-	dread
9)	Certain not to be fatal	-	certain to be fatal

TABLE 4.10

SUMMARY OF THE RESULTS OF THE PRINCIPAL COMPONENT ANALYSIS

PERFORMED BY FISCHHOFF ET AL (1978)

CHARACTERISTIC	LOADINGS		COMMUNALITY
	Factor 1	Factor 2	
Voluntariness	0.89	0.03	0.79
Immediacy	0.70	-0.45	0.69
Known to those exposed	0.88	-0.39	0.93
Known to science	0.88	-0.28	0.86
Control	-0.83	-0.24	0.75
Newness	-0.87	0.14	0.78
Chronic	0.62	0.55	0.69
Common	0.67	0.60	0.81
Severity	0.11	0.91	0.84
Latent root	5.30	1.90	
% variance	58.9	21.1	

As pointed out earlier, one of the tasks performed by participants was an assessment of the perceived risk and of the perceived benefit of the 30 technologies and activities. Fischhoff et al (1978) report that there was no significant correlation between perceived benefit and any of the nine risk characteristics. Perceived risk was correlated significantly ($p < .001$) only with dread and severity.

The researchers plotted each activity and technology as a function of both its perceived benefit and its perceived risk. They then calculated a regression line for the 15 most voluntary activities (as rated by participants) and another regression line for the 15 least voluntary activities. Starr's results suggested that these two lines should be parallel, the line for voluntary activities lying above the other. Fischhoff et al found that the two lines "were virtually identical".

The researchers calculated values of acceptable risk on the basis of both perceived levels of risk and acceptability of these levels. The activities and technologies were re-plotted as a function of both their perceived benefit and their acceptable risk level. This time the regression lines for voluntary and for involuntary activities were different. Furthermore the researchers point out that "for any given level of benefit, greater risk was tolerated if that risk was voluntary, immediate, known precisely, controllable, and familiar".

4.4.2 Methodological limitations

As pointed out earlier Fischhoff et al calculated levels of acceptable risk on the basis of participants' assessment of "current" levels of risk and of their assessment of how much safer activities should be. The researchers point out that such a procedure has a few underlying assumptions in need of verification by future research. The studies by Green and Brown (1977 c, 1977 d) are shedding some light on the validity of these assumptions.

The two factors which Fischhoff et al found and which explain most of the nine risk characteristics have some common characteristics with factors in studies discussed previously. For instance the notions of control and of severity belong to different factors; this was also found by Golant and Burton (1969) and by Champion (1977), and is suggested by the results of E1 and E2 (Green and Brown, 1976 b, 1977 b). However Fischhoff et al used only nine risk characteristics which they selected in advance. Whether the factors would remain the same if other risk characteristics were considered is not known. The results of Fischhoff et al's research indicate which of the nine characteristics are prominent, but they do not make it possible to find out whether other risk characteristics would be considered as more important by the participants.

Fischhoff et al mention that their respondents were "members of the League of Women voters and their spouses". They point out that different types of respondents might express different opinions and perceptions.

4.5 DISCUSSION

The four studies discussed in this chapter were different in many respects. The participants involved, the tasks to be performed, the hazards to be rated and the characteristics on which to rate them varied from one research to the other. Nevertheless themes have emerged from all four which have been the subject of theoretical discussion in Chapter 3. These themes are reviewed briefly hereafter.

4.5.1 Control

Cohen (1970 a), Melinek et al (1973) and Ross (1974) have all made reference, directly or indirectly, to the importance of perceived degree of control over a situation in hazard assessment. In all cases it is argued that a risk is underestimated if a person feels in control of events.

Cohen (1970 a) speculates that criminals' belief in their ability to escape arrest biases their assessment of the risk taken in committing crime. Ross (1974) argues that confidence in their skill is one of the main factors which makes experienced mountaineers assess a climb as less risky than do novices. And Melinek et al (1973) remark that the risk of being in their home when it catches fire is seen as less of a threat by firemen than by other respondents.

The notion of control may be what explains differences found in research on gambling between situations where only chance is perceived to intervene and situations where skill is perceived to have an influence. When it is believed that skill can be used, probabilities of desired outcome tend to be overestimated (Cohen, 1964).

It may also be because of perceived control that people tend to accept higher levels of risk for themselves than they do for those around them (Ross, 1974). This could be because there is a tendency to underestimate other peoples' degree of control over events and possibly to overestimate one's own control.

Some empirical evidence of the importance of the notion of control is present in all four studies discussed above. For instance, "controllability" was the second most important factor in respondents' assessment of hazards, in Golant and Burton's study. Green and Brown (1976 b, 1977 b) mention that degree of voluntariness of an activity is one of the main predictors of perceived safety; they also say that perceived degree of control over risk (potential harm) is one of the main predictors of perceived hazardousness. Finally the notion of "scope for personal control" emerged as the most important factor to be extracted in Champion's (1977) consensus repertory grid. "Control" is also one of the key concepts in Fischhoff et al's first factor.

In the discussion of E1 and E2 it was argued that the concept of voluntariness was a good predictor of perceived safety probably because voluntary activities were deemed to be under control. This suggestion finds some support in Fischhoff et al's results. Their first factor encompasses both concepts of voluntariness and of control. Furthermore the researchers found that the correlation between the two concepts was -0.76 ($p < .001$); thus voluntary activities are seen as controllable.

As pointed out earlier, methodological considerations impose limitations upon the interpretation of the results from all four pieces of research cited above. Nevertheless it may prove significant that the notion of control emerged as important even if three different methods were used to identify it. Such a recurrence may be an indication of the robustness of the concept of control.

4.5.2 Dreadfulness

Another notion has regularly emerged from the literature reviewed in Chapter 3. As postulated by Ross (1974), perceived dreadfulness of potential outcomes of a situation weighs heavily in the assessment of that situation. That is probably why the "magnitude" factor was an important one in Golant and Burton's (1969) study. Possible indicators of dread (e.g. numbers killed, numbers injured) correlated significantly with both perceived safety and perceived hazardousness in Green and Brown's (1976 b) E2 experiment. Champion's (1977) "injury potential" would appear to convey the same underlying notion of dread which is conveyed by "magnitude", "numbers killed", "numbers injured" and other indicators of the same type.

Green and Brown's (1976 a) methodology in the E2 experiment precluded identification of an order of prominence of the various concepts in respondents' assessment of hazards. But such an order of prominence is implicit in principal components as extracted by Golant and Burton (1969), by Champion (1977) and by Fischhoff et al (1978). In all three studies perceived severity of consequences proved to explain less common variance than did perceived control. Furthermore constructs related to severity were mentioned somewhat less often than constructs related to control in E1 (Green and Brown, 1977 b). Therefore it would seem that, although perceived dreadfulness is an important part of hazard assessment, it is not as prominent as perceived control.

4.5.3 Likelihood

In Chapter 3, assessment of probabilities was pointed out as being a part of hazard assessment. A great deal of attention has been devoted, in the literature on hazards, to assessment of likelihood. There were, however, indications that perceived likelihood may be only a minor aspect of hazard assessment.

In so far as the studies reviewed in this chapter are typical, they tend to confirm that perceived likelihood is less important than perceived control and perceived dreadfulness. For instance, if the assumption is made that Golant and Burton's (1969) "expectancy" factor could also be labelled "likelihood", then "controllability" and "magnitude" are more prominent than perceived likelihood. In Green and Brown's (1976 a) E2 results, both perceived safety and perceived hazardousness had overtones of perceived likelihood; but this last concept did not stand out on its own.

It is possible that the relative unimportance of perceived likelihood makes it a difficult factor to extract from respondents' assessments of hazards. This could be why at first sight none of Champion's (1977) factors appears directly related to perceived likelihood. Relative unimportance would also explain why, in Green and Brown's (1976 a) results, perceived likelihood is not identified as a prominent concept.

4.5.4 Other factors

In Golant and Burton's (1969) study the underlying trend which explained the largest portion of variance was rather ambiguous; close scrutiny of this factor revealed that it was very difficult to label it accuretaly. Two factors were also identified in Champion's (1977) results where some ambiguity prevailed.

Principal component analysis was used in Golant and Burton's (1969), Champion's (1977) and Fischhoff et al's (1978) studies. This type of factor analysis generates as many theoretical components as there are items in a test; but it is argued in Chapter 6 of this thesis that determining which of these components are significant can pose a problem.

Slater (1972) argues that, because people perceive the world in three dimensions, they also tend to assess concepts, ideas or even objects along three dimensions. Therefore he postulates that, when principal component analysis is used on a semantic differential test or on a repertory grid, it is likely that three significant components emerge.

It is then surprising that Golant and Burton (1969) and Champion (1977) found four factors whereas Fischhoff et al (1978) found only two. The latter also point out that "the communalities were high, indicating that this two-factor solution did a good job of representing the ratings for the nine scales". It is possible that Fischhoff et al extracted fewer factors because fewer scales were used. In any case in none of these three studies did respondents use scales which they had elicited themselves.

It seems possible that, if respondents were given t opportunity of mentioning risk characteristics which they consider to be most important, the rating of hazards on these characteristics would be meaningful to them. It also seems likely that, as suggested by Slater, three main dimensions would be identified.

4.6 HYPOTHESES

In the light of theoretical considerations raised in Chapter 3 and of empirical evidence reviewed in this chapter, it is possible to present a first formulation of the main hypotheses of this thesis, at least in their broad outlines.

Firstly it is hypothesized that perceived control is the most prominent aspect of hazard assessment. It is hypothesized that perceived dreadfulness is the second most important aspect of hazard assessment. Perceived likelihood, it is postulated, will emerge as the third important aspect.

Furthermore these three facets of hazard assessment are postulated to be prominent for most people and for most hazards. Finally, if other aspects of hazard assessment are identified they will prove to be of less importance than those mentioned above.

This is only a first formulation of the hypotheses of this thesis. A second formulation is proposed, in operational terms, at the end of Chapter 5 in view of the methodological considerations which are discussed in that chapter.

All studies discussed in this chapter have identified important aspects of hazard assessment. Some factors were different, others had many characteristics in common. It is as though, because of different methodologies, different pieces of a puzzle have been located. It is hoped that, by allowing respondents to express what risk characteristics they consider to be important, all pieces of the puzzle can be assembled to form a coherent picture of the way in which people assess hazards.

CHAPTER 5 : METHODOLOGY

Once hypotheses are formulated, means must be sought for testing them. Amongst other things decisions have to be made about the choice of research techniques and analysis of results.

Some comments were made in the previous chapter about the methodology of other studies. More specific consideration will be given in this chapter to the measurement and analysis techniques used in these studies. The examination of other studies suggests that the repertory grid technique is the most appropriate instrument for this research. Therefore, more detailed attention is paid in this chapter to the grid technique and to the theory of personal constructs from which it stems. After appropriate techniques for the analysis of repertory grids are identified the hypotheses of this research are redefined in operational terms.

5.1 CHOICE OF A MEASUREMENT INSTRUMENT

5.1.1 Previous research

One of the noticeable features of the four studies reviewed in Chapter 4 is that despite specific orientations, their general aims show similarities. For instance, in all four studies, instruments were devised to identify the meaning which people attach to hazards. To a certain extent, it can also be said that these instruments measured peoples' perceptions of hazards.

Therefore, it is not surprising that these four instruments also show similarities. All four instruments were made up from the same two types of components. Firstly, they involved a list of hazards which were rated by respondents. Secondly, all four included a series of scales which respondents used in order to rate the hazards.

So, the general format of the instruments was similar in all four studies. There were however some important differences between them. Differences between instruments stemmed mainly from the way in which these were assembled.

For instance, Golant and Burton (1969) used a semantic differential test for which hazards and scales were provided by the researchers. Green and Brown (1976a) used a repertory grid for which the researchers supplied the list of hazards; respondents themselves provi[illegible] the criteria which they subsequently used to rate the hazards. Champion (1977) also used a form of repertory grid; hazards were spotted by respondents in their working environment, then criteria or scales were elicited during a series of pilot studies and supplied to the original respondents in order for them to rate the hazards on these scales. Finally, Fischhoff et al (1978) selected the rating scales ("risk characteristics") as well as the elements ("activities and technologies") which constituted the measurement instrument. Thus, whilst two studies used a semantic differential test or the equivalent, two other studies used some form of repertory grid.

Green's (1975b) arguments in choosing the repertory grid were quoted in Chapter 4. Considering these arguments and the weaknesses of Golant and Burton's (1969) questionnaire discussed in the previous chapter, it seems logical that the repertory grid should be chosen. However, it is important to fully understand the repertory grid as compared to other instruments. Some of the finer points find their roots in the personal construct theory from which the repertory grid was developed.

5.1.2 Theoretical considerations

When he proposed the repertory grid, Kelly (1955) stated that his technique stemmed from the principle that "man looks at his world through transparent patterns or templets which he creates and then attempts to fit over the realities of which the world is composed". The four studies reviewed in Chapter 4 all attempted to identify which of these "transparent patterns or templets" were relevant to hazard assessment. Practically all attitude and opinion measurement techniques are concerned with the identification of such templates (Edwards, 1957).

There are, however, important differences in other assumptions underlying the various instruments. Kelly (1955) expresses these differences in this way:

> "There are two ways in which one can look at psychological measurement and clinical diagnosis. On the one hand, he can seek to fix the position of the subject with respect to certain dimensions... or to classify him as a clinical type... On the other hand, he can concern himself with the subject's freedom of movement, his potentialities, the resources which can be mobilized, and what is to become of him".

The way in which these considerations are formulated reflects the clinical setting from which they evolved. But Bannister and Mair (1968) argue that the basic principles underlying these considerations also apply to other forms of assessment outside the clinical setting.

Kelly (1955) argues that the first alternative described above is tantamount to assessing a respondent by using templates which may not be relevant in the respondent's assessment of reality. These templates may have been chosen by a very rigorous scientific process, and yet the risk remains that for a respondent these templates may be peripheral or secondary. Kelly (1955) goes on to argue that the second alternative corresponds to assessing a respondent by identifying and analysing the respondent's own way of construing the world. In other words, the first approach would represent an attempt to categorize a respondent according to the way in which the respondent construes events on a set of fixed criteria. The second alternative refers to an elicitation of the criteria which the respondent uses to construe the events.

These approaches, according to Kelly, have important clinical implications. For instance, fixed criteria tend to yield a descriptive representation of a person, whereas a dynamic analysis can usually be achieved by using a respondent's own constructs. Furthermore, criteria used by clinicians tend to evolve slowly and to be an outdated standard of assessment. On the other hand, as respondents evolve so do their representations and assessments of the world. Identifying those representations, comparing them between respondents and monitoring their change over time may prove to be a clinician's best way of understanding respondents' behaviours and of predicting their evolution.

5.1.3 Personal construct theory

It is on the basis of such considerations that Kelly (1955) developed the personal construct theory. It would be beyond the scope of this thesis to review Kelly's theory extensively. However, some of its basic concepts have direct implications for the choice of a measurement technique.

Therefore, the main foundations of the theory are listed in this section for the purpose of explaining the essence of Kelly's arguments. In order not to embark upon arguments which are not directly relevant to this thesis, no definition of terms is undertaken other than that which is provided in Kelly's basic statements. However, those statements which are relevant are discussed further.

Personal construct theory rests on what Kelly (1955) describes as a fundamental postulate and eleven corollaries. The fundamental postulate is formulated in this way: "A person's processes are psychologically channelized by the ways in which he anticipates events" (Kelly, 1955).

This postulate has a number of underlying assumptions. In turn, the postulate can be expanded into a number of connotations. These assumptions and connotations are what Kelly calls corollaries. Below is a list of the eleven corollaries, along with their titles, as formulated by Kelly.

1) A person anticipates events by construing their replications (construction corollary).

2) Persons differ from each other in their constructions of events (individuality corollary).

3) Each person characteristically evolves, for his convenience in anticipating events, a construction system embracing ordinal relationships between constructs (organization corollary).

4) A person's construction system is composed of a finite number of dichotomous constructs (dichotomy corollary).

5) A person chooses for himself that alternative in a dichotomized construct through which he anticipates the greater possibility for extension and definition of his system (choice corollary).

6) A construct is convenient for the anticipation of a finite range of events (range corollary).

7) A person's construction system varies as he successively construes the replications of events (experience corollary).

8) The variation in a person's construction system is limited by the permeability of the constructs within whose ranges of convenience the variants lie (modulation corollary).

9) A person may successively employ a variety of construction subsystems which are inferentially incompatible with each other (fragmentation corollary).

10) To the extent that one person employs a construction of experience which is similar to that employed by another, his psychological processes are similar to those of the other person (commonality corollary)

11) To the extent that one person construes the construction processes of another, he may play a role in a social process involving the other person (sociality corollary).

It is difficult to summarize Kelly's theory more succinctly; these corollaries are summaries in themselves. Further attempts at rephrasing would run the risk of major oversimplification.

However, some aspects of personal construct theory suggest important considerations in the choice of a technique for assessing respondents' perceptions of the world. For instance, the organization corollary implies that people devise their own method of assessing events; the individuality corollary states that methods of assessment vary from person to person. The basic postulate of the theory suggests that understanding a person's system of assessment of events is the key to understanding the person.

It would appear from the experience and fragmentation corollaries that a person's assessment system is arrived at by trial and error, and that it may evolve over time and with every new event. The theory also suggests that a person uses different assessment systems for different types of events; therefore, an assessment system is dependent upon the nature of assessed events. To a certain extent, the nature of events would also determine what constructs can be used to assess them.

Some of these corollaries have direct implications in statistical terms. For instance, statisticians take the individuality corollary for granted in their notion of "individual differences" (Guilford and Fruchter, 1973). In section 4.1.3 of the previous chapter, where Golant and Burton's semantic differential test is discussed, some preliminary indications are given of other implications of these corollaries. These implications are analysed in detail in this next section.

5.1.4 Limitations of the semantic differential test

In view of the above discussion some weaknesses of a semantic differential test such as the one used by Golant and Burton (1969) become apparent. For instance, by supplying scales along with a list of hazards to their respondents, Golant and Burton sought to locate respondents on a set of fixed criteria rather than to identify the respondents' own construction systems. In so doing, it was argued, they may well have limited the scope of their findings to those concepts which they had built into their test. Therefore, the factors which they identified may not have been those of greatest importance or relevance to respondents. The fact that the four significant factors explained between them less than half of the variance of the test may be an indication that important concepts were overlooked.

Supplying rating scales to respondents may create another type of problem. As pointed out earlier, Kelly (1955) states that "a construct is convenient for the anticipation of a finite range of events only". This range of events, outside which a construct becomes ambiguous, meaningless or redundant, is known as "the range of convenience" (Kelly, 1955) of that construct.

When respondents have to rate items outside the range of convenience of the scales they use, then their rating may be ambiguous and difficult to interpret. Bannister and Mair (1968) note that: "This sort of objection has been made on general grounds by Brown (1958) in a review paper significantly entitled "Is a boulder sweet or sour?" ". They also mention that "One of the present authors has for years, in papers and lectures, used the attempt to designate false teeth as either religious or atheist as an example of possible difficulties over range of convenience" (Bannister and Mair, 1968). It was argued in Chapter 4 that ambiguities of this type have imposed limitations on the usefulness of Goland and Burton's (1969) results.

Because of those "possible difficulties over range of convenience", Bannister and Mair (1968) argue that the semantic differential test can prove inadequate in measuring the "meaning" attached to certain concepts or elements. It was to obviate those difficulties that the repertory grid technique was developed from the personal construct theory.

5.1.5 Beginnings of the repertory grid

Originally, the repertory grid was devised as "a new diagnostic instrument which illustrates how (personal construct theory) can be applied to the practical needs of the psychotherapist" (Kelly, 1955). The grid was first used to analyse respondents perceptions of the persons of importance to them and the relationships between respondents and those persons. A sociogram is used to depict a network of relationships among a group of individuals; the original purpose of the grid was to analyse the dynamics of one respondent's relationships with others.

In its first version, the grid consisted of a series of 24 role descriptions. These could be presented either verbally, or as a printed list, or separately on cards. Respondents were asked to name, for each description, a person whom they knew and who corresponded to that description.

Then, descriptions (and names) were presented again, three at a time, to the respondent. The respondent was asked to explain how any two of the persons represented by the descriptions and named earlier were similar to each other and different from a third.

The grid could be administered using recommended set groups of three descriptions. It could also be done by using the "sequential version", i.e. presenting descriptions numbers 24, 23 and 22, then 23, 22 and 21, and so on.

The similarities and differences mentioned by the respondent were considered as criteria or "constructs" which the respondent used to assess those persons represented by the descriptions. Then, the respondent rated those persons on each construct.

Rating was done in one of two ways. For instance, each person on the list could be examined against each construct and assigned one of the labels of the construct (e.g. warm or cold). Alternatively, the respondent could sort the persons into two groups, each group being assigned one of the labels; sorting was carried out for each construct.

Kelly (1955) listed eight variations of the above procedure, including a version which could be administered to a group rather than to individuals. All these versions dealt with roles and relationships, and were primarily for clinical use. Kelly (1955) suggested mathematical analyses which could be performed on grids; he did not, however, elaborate on other uses to which the grid method could be put.

5.1.6 Developments of the grid

Other researchers developed numerous alternative uses for the grid. For instance, Fransella and Bannister (1977) list some 225 references dealing in whole or in part with the repertory grid; these references were related to subjects including, among others, the causes of stuttering, the nature of schizophrenia and levels of cognitive complexity. But alternative uses have not been the only development of the repertory grid.

Fransella and Bannister (1977) mention that a more flexible way of eliciting the elements of a grid has been used in some studies. Instead of being provided with a list of items to be rated or assessed, respondents in some experiments were asked to name a series of objects, persons, concepts or whatever which fell into a given category. In such cases, the instruction was of the type: "name some people you like", for example. These then became the elements from which constructs were elicited and which were rated. No study could be traced which compared the reliability of this method to that of providing a list of elements.

The other important development mentioned by Fransella and Bannister (1977) was the introduction of metric techniques for scoring grids. As described in the original grid technique, constructs were treated as dichotomies. Greater flexibility in the use of constructs has been achieved by the use of ranking procedures or of rating scales. It is interesting to note that the instruments (including the repertory grids) used in the four studies reviewed in the previous chapter all included rating scales. The use of rating scales appears to be in contradiction with the dichotomy corollary. However, Bannister and

Mair (1968) argue that if a construct is genuinely dichotomous, only two points of the rating scale - usually the extremes - are used. They also mention that some judgements are better expressed in terms of degrees rather than as dichotomies. Fransella and Bannister (1977) point out that the format of a rated grid is very similar to that of a semantic differential test; but they insist that there are still important differences between the two techniques.

5.1.7 Choosing the repertory grid

As pointed out earlier, according to Bannister and Mair (1968), semantic differential tests impose a rigid assessment framework upon respondents. The same general criticism can be directed at most attitude and measurement techniques (Edwards, 1957). It is not sure whether all constructs which are relevant, from a respondent's point of view, are present in a test such as the semantic differential. Furthermore, the possible absence of relevant constructs makes it difficult to find out if those constructs which are supplied are the most relevant.

Fransella and Bannister argue that, since repertory grids do not impose upon respondents as rigid a framework as do other techniques, the important and relevant dimensions are bound to be elicited. They go on to argue that a more flexible framework somewhat reduces experimenter bias. Other types of experimenter bias are not removed. For instance, it is assumed, in the grid technique as well as in other techniques, that the scale labels have the same meaning for the experimenter and for the respondent. When principal components are labelled, experimenter bias may influence the choice of labels regardless of the type of instruments from which the components were extracted. A more detailed discussion of the assumptions underlying the grid is found in the next section of this chapter.

It was because they did not want to impose a rigid framework that both Green (1975b) and Champion (1977) opted for a repertory grid as their main research instrument. Both researchers devised their constructs as Likert-type five-point scales. Rating scales were prefered to ranking procedures on the grounds that ranking a large number of elements (21 and 19 respectively) could prove to be an awkward task for respondents. Dichotomous constructs were also discarded because of the low level of discrimination inherent in such constructs.

In view of these considerations, a rating-scale grid seemed an appropriate choice for this study.

5.1.8 Assumptions underlying a grid

When a researcher decides to use a repertory grid, he is making certain assumptions. For instance, as pointed out earlier, it is assumed that those dimensions which are important and relevant from the respondents' point of view are elicited and used in the grid. Kelly (1955) describes six other assumptions underlying the use of a repertory grid.

Firstly, he mentions that, when constructs are elicited, it is taken for granted that they can also be useful in assessing elements other than those from which they were elicited. All elements of a grid are usually rated on all constructs. It is therefore assumed that a construct elicited from a triad of elements is also useable in conjunction with the other elements of a grid.

The second assumption is that of the pre-existence of constructs. In other words, those constructs which are elicited are assumed to have evolved from respondents' experiences rather than to have been "concocted on the spot" (Kelly, 1955).

A third assumption is that elements in a grid are representative. That is to say, it is assumed that elements of a grid are chosen so as to elicit all possible components of a construct system.

Fourthly, it is assumed that constructs which are elicited are related to respondents' understanding of the elements of a grid. This assumption is closely related to the fifth one listed by Kelly. This next assumption is that the construction system which emerges from a grid is the one which a respondent uses as a basis for his behaviour. In other words, it is taken for granted that a respondent does in practice use that construction system which is identified through his grid.

Kelly refers to his sixth assumption as "that of the functional communicability of the constructs elicited". In other words, it is assumed that the way in which a respondent names a construct enables an examiner to get "some practical understanding" of the meaning attached by the respondent to the construct.

This thesis is not directly concerned with the prediction of behaviour. Thus, the assumption that the construct system which is identified is actually used as a basis for behaviour is not central to the argument of this thesis. Other assumptions, however, may have an important bearing on the interpretation of the results of this research. Therefore, more detailed attention must be paid to these assumptions.

5.1.9 Control of the assumptions

It was not the main object of this thesis to test the accuracy of these assumptions. On the other hand, it was thought important to make sure that the hazard assessment mechanisms identified by repertory grids stood good chances of being genuine. Thus, methodological procedures were adopted in order either to override or to verify some of the assumptions.

Little could be done to make sure that only permeable constructs would be used. However, it was decided to verify the extent to which those constructs which were used were permeable. Thus, giving respondents the opportunity of indicating, by a zero for example, that a construct is irrelevant to the assessment of an element might provide clues about the permeability of that construct. As pointed out in Chapter 6, permission was obtained to analyse grids collected in EI and in which naughts had been used.

Another procedure was used in an attempt to obtain indications about the permeability of constructs. On the basis that the more permeable constructs should be useable in conjunction with a wider variety of elements, it was decided to use four different sets of hazards as elements. The more permeable constructs should emerge as relevant to all types of hazards.

Using different sets of hazards was thought to have another advantage. As pointed out in the previous section, the use of a repertory grid rests on the assumption that the elements used are representative. It was thought that if the assessment patterns were found to be similar for all sets of hazards, then to a certain extent, it could be said that any of the four sets of hazards is representative of the others.

The assumption of "functional communicability of constructs" is common to practically all measurement techniques. For instance, the use of a semantic differential rests on the assumption that the respondent understands the rating scales supplied to him in the same way as the experimenter did when devising these scales. In the repertory grid the same assumption applies, but the other way around; in other words, it is assumed that the experimenter understands a construct in the same way as did the respondent who coined it. It was thought that, if constructs were discussed, their meaning would become more obvious and the chances of erroneous interpretation would be reduced. One method which came to mind was a group discussion. Other reasons, described in subsequent paragraphs of this section, could be put forward to justify the use of a group discussion. It is mentioned in Chapter 6 that such a discussion took place in the second sample as the way to assemble their grid. In the course of this discussion, some clarification of meaning did indeed take place; it was, however, the exception rather than the rule. Thus, it appeared that most constructs were relatively straightforward and needed little or no further explanation. This was also an indication that most respondents understood the constructs in the same way.

Another advantage of the group discussion became obvious while the discussion was taking place. Although it is difficult to quantify such an observation, a cross-fertilisation phenomenon was observed. In other words, a construct mentioned by a respondent would remind other respondents of other constructs which they would immediately mention. Possibly, these constructs elicited as the result of an association of ideas might have been mentioned anyway, in individual interviews, in conjunction with other triads. It seems likely, however, that some salient constructs might have been overlooked if a group discussion had not taken place. Thus, the assumption that a repertory grid ensures the elicitation of salient or important concepts may not be true in all cases.

The choice of an appropriate method of grid analysis can ensure adequate assessment of the relative importance of constructs. It was argued earlier that, because repertory grids allowed respondents to use constructs which they themselves supplied, grids stood better chances than did other instruments of identifying assessment patterns representative of respondents' own preoccupations. However, no method could be devised for obtaining data about the assumption of preexistence of constructs without changing this thesis into a grid validation study. Other studies could be devised, using interview techniques along with repertory grids for example, to test the assumptions.

Kelly's fourth and fifth assumptions reflect the originally clinical use of the repertory grid. As a basis upon which to devise a counseling or therapeutic strategy tuned to a client's needs, and as a tool for assessing the efficiency of that strategy, it had to be assumed that the repertory grid identified assessment patterns which were the keys to a person's behaviour. However, it had been decided that this thesis should concentrate on the identification and exploration of the nature of hazard assessment patterns rather than on their behavioural implications. Thus, Kelly's fourth and fifth assumptions were not looked into in this thesis. It might prove interesting and useful, however, to develop further research to look into influences upon hazard assessment patterns and behavioural implications of these patterns. Suggestions for such research are mentioned in Chapter 12.

Some methods (use of zero ratings, group discussion, different series of elements) were used to override or control some assumptions. It is possible that these techniques influence or alter the assumptions, or even introduce new assumptions of their own. For instance, the use of a group discussion rests on the additional assumption that being in a group does not prevent a respondent from mentioning certain constructs. Furthermore, it is not known whether discussing a construct with a group of respondents affects its permeability. Again, further research on these specific points would be needed.

The very nature of the chosen instrument and the specific purpose of this research suggested grid analysis strategies and techniques. The next sections of this chapter describe the various analysis techniques chosen and the reasons for their choice.

5.2 ANALYSIS OF INDIVIDUAL GRIDS

The main aim of this project was to test whether there are hazard assessment patterns which are common to most individuals and which are applicable to most types of hazards. Therefore, methods of analysis had to be devised in order firstly to identify hazard assessment patterns from individual grids, secondly to compare those patterns between respondents for each type of hazards, and thirdly to compare patterns across the various types of hazards.

Each of these three steps involved methodological problems of its own. Therefore, each step is dealt with separately. In this section, a method is discussed for the first step, i.e. the analysis of individual grids.

5.2.1 Processing of individual grids

Ingrid 72 is a fortran program devised by Slater (1972) for the analysis of individual repertory grids. It provides a number of important statistical calculations. Firstly, it estimates the variance of each construct; it is possible to compare this variance with a measure of bias (deviation of the means from the mid-point of their respective scales) in order to examine the way in which the rating scales were used.

There are instances where scales are not used evenly between constructs. The program offers the option of normalising the scales for the remaining analyses, i.e. rescaling them "so that they each have their total variation put equal to unity" (Slater, 1972). While not altering the relationship between the ratings on a construct, normalisation makes construct scales comparable. Failure to request normalisation can lead to very peculiar results in principal component analysis.

The program then rejects those constructs which show no variation about their mean; in other words, constructs on which all elements have received an identical rating are omitted from the analysis. Then, the variation of each element across the constructs is listed. One of the most important features of the analysis is listed next: the matrix of correlations and angular distances between constructs.

Whereas correlation coefficients are not additive, angles are. Since correlation coefficients can be calculated from angular distances, the latter could be useful for calculating mean correlations. As for the correlations, although principal component analysis will indirectly be based on them, they are put to no other use by Ingrid 72. As described later, however, further calculations based on the matrix of correlations were performed.

Grids are usually analysed as a function of their constructs. In some instances, however, information about elements can be very useful. For example, elements which show no variation across constructs would most likely require a different set of constructs altogether in order to be properly assessed. As constructs are correlated on the basis of their score on the same elements, elements showing no variation in fact introduce biases in the correlations. That is why the next listings of Ingrid 72 are the variance of each element and the distances between elements.

Then, if requested in the options of the program, the matrix of "Sums of products" is printed. "It is of no direct interest for interpreting the grid" (Slater, 1972), but it is nevertheless an important feature as "it is the matrix to which the principal component analysis is actually applied" (Slater, 1972). The main feature of Ingrid 72, the principal component analysis, is listed next by the program.

First, the latent root of each component is printed. Then the Bartlett test is performed in order to work out how many components are significant. Then, the grid is printed again; but raw scores for elements are converted into deviations from their construct means. This transformed grid is followed by specifications of components.

Slater (1972) says that principal component analysis was selected because it made a component analysis possible for elements as well as for constructs. If axes were to be rotated in order to obtain a better dispersion of elements, then construct specifications would become meaningless. Similarly, a better dispersion of constructs by rotation would render element specifications useless. For that reason Slater (1972) did not include an algorithm to rotate axes in Ingrid 72.

Each component is therefore listed by Ingrid 72 as a function of the elements as well as of the constructs. Elements and constructs are described on each component as a function of their vector, their loading and their residual (or unexplained) variance. Components are listed three at a time.

After the first group of three components, polar coordinates for elements and constructs are calculated and listed. The polar coordinates can be plotted on a two-dimensional graph; the cartesian axes of that graph can be viewed as two psychological dimensions.

Following the specifications of each group of three components, and following the polar coordinates after the first group of three, matrices of residual deviations are printed. One matrix is printed for each component which has been extracted. Each cell of a matrix is the variation of an element on a construct which remains to be explained, once the variance explained by the components extracted so far has been subtracted.

The number of components which are listed in full is dependent upon the option card of Ingrid 72 and upon the result of the Bartlett test. A user of Ingrid 72 must specify the minimum and maximum number of components to be listed in full. The minimum number will be compared to the number of components found significant by the Bartlett test. Polar coordinates are printed only if the full specifications of at least three components are printed. The greater of the two values will become the minimum number of components which will be listed. The maximum number specified by the user will be compared to the number of significant components. The smaller value will become the maximum number of components to be listed in full.

Ingrid 72 is stopped after the maximum number of components have been listed with the associated matrices of residual deviations. When the results of all these calculations become available, they represent a sizeable amount of information to be interpreted.

Impressive though they may look, these statistics and calculations may be irrelevant, inadequate or insufficient for the purpose of this research. At this stage, it seems appropriate to examine the relevance of these statistics to the aim of this thesis.

5.2.2 Data reduction methods

Kelly (1955) argues that a person's way of construing an object occupies a certain area of that person's mind. When either similar or related, or even categories of objects enter into consideration, then the corresponding area of the person's mind expands and becomes increasingly complex; it evolves into a "psychological space" with a structure of its own.

Constructs can be viewed as dimensions of this psychological space. Kelly points out, however, that these dimensions are not all mutually independent. Some inter-relatedness between constructs can usually be observed. Such inter-relatedness may mean, for instance, that some constructs are facets, connotations, explanations or sub-sections of other constructs. By analogy, Kelly uses the word "context" to describe the inter-relatedness of constructs. He argues that studying contexts "may give us an understanding of the inter-weaving of a client's terminology and provide us with an understanding of his outlook which no dictionary could offer". (Kelly, 1955).

Thus, all constructs do not have the same status. Some constructs constitute the context of other, more basic dimensions of the psychological space. To a certain extent, there are hierarchical relationships among constructs. A person's construing system can be summarized by the few central constructs which account for or influence most of the others. Slater (1972) argues that, as we perceive the physical world in three dimensions, we also tend to develop three-dimensional construing systems. Kelly (1955) refers to studies which indicate that repertory test "protocols tended to be factorially simple, in some cases reducible to two factors".

In a clinical setting, identifying which constructs are more central or basic may prove to be very worthwhile. As these central constructs are the cornerstones of a person's construing system, they should be the focus of a clinician's attention. Kelly argues that any therapeutic or counselling intervention tuned to these central dimensions stands good chances of reaching a client.

Research results quoted by Kelly (1955) indicate "that the number of constructs produced by a subject tends to be quite limited". However, the task of identifying the underlying structure can be cumbersome. Even a limited number of constructs can represent a large number of inter-relationships whose structure is not always accessible to conscious thought or logical analysis. Asking a client to become aware of the intricacies of his own construing system is likely to be very time-consuming. Kelly prefered to turn to statistical ways of finding out how constructs are structured. Therefore, he devised a non-parametric factor analysis which extracts the main dimensions underlying dichotomous constructs.

The studies by Golant and Burton (1969), Champion (1977) and Fischhoff et al (1978) all included some form of factor analysis for the identification of a structure underlying hazard assessment criteria. The lists of criteria (or constructs) they studied were different, thus the structures they identified were also different.

Some similarities between these structures were pointed out in Chapter 4. These similarities can be interpreted as indications that these studies identified parts or facets of the general structure of hazard assessment.

The four studies reviewed in Chapter 4 represent an impressive sum of criteria or constructs. The main hypothesis of this research is that these and other constructs are basically arranged in most respondents' minds into a three-dimensional structure; in other words hazard assessment, it is hypothesized, can be summarized into three basic dimensions underlying or representing most constructs; the parts of this three-dimensional structure which other studies identified depended on their selection of constructs or rating scales.

Thus, it was necessary to choose a statistical instrument which made it possible to identify such an underlying structure. More specifically, the chosen statist'cal technique would have to perform three tasks. Firstly, the interactions between constructs had to be taken into account. Secondly, as Rump (1975) describes it, on the basis of these interactions, a large number of ratings and scores had to be reduced to a smaller, more manageable number of underlying trends. Thirdly, it was expected that the statistical instrument would provide some indication of the relative importance of the underlying trends. These three task characteristics are common to most statistical instruments known as data reduction methods.

The various forms of factor analyses are probably the most widely used category of data reduction methods. For the analysis of repertory grids, Kelly (1955) and Slater (1972) advocate the use of one form or another of factor analysis. As pointed out earlier, factor analyses were chosen by Golant and Burton (1969), Champion (1977) and Fischhoff et al (1978). More specifically, Slater (1972) argues that principal component analysis without axis rotation is the most appropriate statistical tool for the analysis of repertory grids. Champion (1977) and Fischhoff et al (1978) adopted this specific technique; the latter further point out that the axis rotation which they performed brought practically no additional precision in dimension specifications. Therefore, principal component analysis without axis rotation was retained for testing the hypotheses of this research.

However, relying solely upon principal component analysis, especially since no rotation solution was to be used, seemed a doubtful strategy for two reasons. Firstly, as pointed out at the end of this chapter, it is hypothesized that identified trends are similar between respondents and between samples. Unless reliable or robust trends are compared, trend differences might be found which could in fact be attributable to weaknesses in the statistical technique. Child (1976) suggests that, in order to test the reliability of trends, it might be adviseable to use more than one type of factor analysis on each set of data.

Secondly, there are other types of data reduction methods. Besides factor analyses, the other widely used type of data reduction methods encompasses cluster analyses. Both Lesage (1973) and Rump (1975) argue that there are grounds for adopting cluster analysis rather than factor analysis. These authors state that cluster analyses are more stringent, that their results are more robust and that the format of these results facilitates a more detailed interpretation.

Principal component analysis is based on correlations between items. In using this analysis, it is assumed that all common variance between items is due to a certain similarity in concepts between the items. The analysis attempts to apportion all the common variance into various similarity trends. Finally, all items are compared to all trends; any item can be associated with more than one major trend.

Like principal component analysis, most cluster analyses are based upon correlations between items. In cluster analysis, however, no similarity trends are estimated. The assumption underlying these methods is that those items which have a concept in common all correlate significantly among themselves. Therefore, unlike factor or component analytical methods, most cluster analytical methods presume only one major underlying trend in each item.

It is quite apparent that using both types of methods involves m 'ing different types of assumptions. No a priori reason could be identified why either type of assumption would be more likely in this research. Therefore, it was decided to perform a cluster analysis and a principal component analysis on each set of data. Such a procedure should facilitate the identification of cases where either type of assumptions is erroneous. Furthermore, this strategy should provide indications about the robustness of underlying trends; trends identified by both analyses are very unlikely to be statistical artefacts.

5.2.3 Use of Ingrid 72

Ingrid 72 was used mainly for its principal component analysis. When each grid was processed, full listing for a minimum of 3 and a maximum of 10 components was reauested. In some grids, more than 10 components were found to be significant by the Bartlett test. In these cases, the grids were re-processed with a request for a maximum number of fully listed components equal to the number found significant in the previous processing.

Ingrid 72 was found to be advantageous over other principal component analysis packages for three reasons. Firstly, Ingrid 72 was devised specifically for the analysis of repertory grids; therefore, it provides specifications of both elements and constructs for each component. This feature proved useful for interpreting those components which were found to be significant but which explaines a relatively small proportion of the variance (15% or less). Although this research concentrated mostly on constructs in the grids, some components were labeled only after scrutiny of the elements which loaded significantly, whether positively or negatively.

Secondly, Ingrid 72 provided information about statistical characteristics of constructs and elements (e.g. variance, bias, etc..) This information also proved useful in interpreting the results. But most important, Ingrid 72 also provided the matrix of correlation coefficients which made it possible to perform the chosen cluster analysis.

The details of the two major analyses as used in this research are discussed in the following sections.

5.2.4 Principal component analysis

Principal component analysis is a statistical method which attempts to isolate the most important concepts or trends underlying a test or a series of tests. The aim of the method is also to estimate the magnitude or the importance of those trends. The magnitude of a trend is estimated by the proportion of the total common variance of the test (or tests) which can be accounted for by the trend. Magnitude is expressed as the sum of the squared distances between the items of the test and the component (trend); it is generally referred to as "sum of squares" or as "latent root".

But an absolute value of a latent root can be somewhat meaningless. Therefore, the sum of squares for a factor is compared to the total variance of the test, i.e. to the sum of sums of squares; this ratio is expressed as a percentage. Thus, a component with a latent root of 0.92 might explain 50% of the variance of a small test but only 27% of the variance of a larger test; in the latter case, the trend would not be as important as in the former.

Some advantages of this technique in the analysis of individual grids have been mentioned by Slater (1972) and quoted in the previous section. Salter also discusses why this method is preferable to other types of factor analysis. It seems that this statistical technique can be very useful for investigating the hypotheses of this research. These hypotheses, briefly pointed out at the end of Chapter 4, and formulated more accurately at the end of this chapter, postulate that there are three main trends in peoples' assessment of hazards. It was argued earlier that the methodological problem is therefore one of extracting underlying trends from the grids. And it appears that principal component analysis is suited for that purpose.

5.2.5 Testing for significant components

Assuming that principal component analysis is used, this poses the problem of deciding when trends cease to be important ones. In other words, ways have to be found for deciding when components are no longer significant. There are a number of methods for making such a decision. All these methods assume that components are listed in order of magnitude, component 1 having the greatest latent root and component N the smallest root.

Two methods are described by Child (1976) as being the most widely used. They are Cattell's Scree test, and Kaiser's criterion of a latent root of 1. In addition, Ingrid 72 provides the results of the Bartlett test.

Cattell argues that the latent root of the various components can be plotted as a function of their rank. The resulting curve should look like a cliff, the size of the latent roots falling abruptly with the first components then levelling to a near-horizontal slope. The method consists of accepting as the last significant factor the first one at the beginning of the near-horizontal portion of the curve.

Kaiser (1959) devised a criterion which is slightly easier to apply. All that is involved is to reject those factors which have a latent root of less than 1. But, as pointed out earlier, a given value for a latent root can have different levels of significance in tests of different sizes. Therefore, this method can be somewhat misleading.

The main characteristic of principal component analysis is that it assumes as many theoretical components as there are items in a test. But most methods for testing the significance of components rest on the assumption that some components explain a negligible portion of common variance. Furthermore, it is assumed that there are practically no differences in the size of those components.

The Bartlett test, described by Slater (1972) and part of Ingrid 72, was devised on the basis of these assumptions. A chi-square test is applied at first on the two components which have the lowest latent roots. If the chi-square is not significant, it means that there is no difference between those components and that the variance left over from other more important components is distributed randomly between all the components being tested. The test is repeated by including the component with the next lowest latent root, until a chi-square test results in a significant value. All those components between which residual variance is distributed at random are rejected and the others are retained as significant.

Both the Bartlett test and Cattell's Scree test postulate a marked difference between the latent root of the significant and of the non-significant components. Both tests expect a discontinuity or a drop in the values of the latent roots. Kaiser's criterion makes the further assumption that the discontinuity occurs at a latent root value of 1. The more assumptions are made about changes in the level of significance, the less reliable the criterion becomes. For this reason, Kaiser's criterion was not used as the main indicator of the number of significant components.

Cattell's Scree test was devised from a large number of factor analyses for which latent roots were plotted. Cattell (1952) points out that in a vast majority of analyses the first component on the near-horizontal portion of the curve was the last component whose meaning could be interpreted. Thus, it could be argued that the Scree test is based on observation rather than on mathematical rationale. Nevertheless, it seems that the Scree test is the visual equivalent of the mathematics of the Bartlett test.

Arguments for preferring either test over the other were not overwhelming. Child (1976) points out that no method for determining significant components is infallible. In the end, experience dictated a line of conduct. It proved nearly impossible to interpret certain Scree test graphs because there were more than one marked changes of slope. However, Ingrid 72 always provided a Bartlett test result. Thus, it was decided to retain the Bartlett test as the main indicator of the number of significant components. In those cases when, as described below, the results of the Bartlett test could be questionned, the Scree test was reverted to.

In some instances, the Bartlett test can be misleading. For example, when latent roots show values which diminish very gradually instead of falling abruptly at one point, the test will indicate no significant component. On the other hand, there are instances where some constructs have a very large intrinsic variance. These constructs usually lead to other constructs having a comparatively small variance. In such cases some latent roots become artificially high and others artificially low. The Bartlett test on such latent roots will lead to some components being rejected as non-significant;

these same components might be considered significant if the latent roots were not artificially altered. As mentioned earlier, when there were indications that such problems had occurred the Scree test was performed. Occasionally for comparison purposes Kaiser's criterion was also applied but not relied upon. Some comparative results of the three tests are discussed in Chapters 8 and 9.

5.2.6 Significant members of components

Once the number of significant components has been determined, another problem arises: that of determining which parts of a test are related significantly to the emerging concepts or components. The usual index of relatedness of an item (or test) to a component is known as the loading. Although calculated in a different way, the loading is in fact a correlation coefficient. Loadings can therefore be used to determine which items are and which items are not correlated to a component.

Burt and Banks (1947) point out that the more components or factors that are extracted, the smaller the proportion of the total common variance which each new component explains. As the proportion of unexplained variance decreases, there is an increasing likelihood that what variance is left is either error variance or unique variance. However, as components are meant to explain common variance, the more components that are extracted, the more stringent the criterion must become in order to retain only those items which contribute to common variance.

A formula was devised by Burt and Banks (1947) to calculate a criterion which becomes more rigorous with each successive component extracted. The cut-off point (minimum loading) for the first component is the coefficient of correlation which is significant at a predetermined level of probability for the number of items in the factor analysis. For each component, this correlation coefficient is multiplied by the expression $\sqrt{\frac{n}{n+1-r}}$, where n is the number of items in the analysis and r is the rank of the component being considered. This formula was used for identifying significant members of components.

5.2.7 Auclair's cluster analysis

Lesage (1973) discusses various types of cluster analyses. He argues that there is little which differentiates the various analyses; all of them give very similar, if not identical, results when applied to the same sets of data. Most of them were not readily available; they were part of statistical computer programs which were not among the software of the Aston computer. One of the cluster analytical techniques, however, could be used quite easily. The technique known as Auclair's cluster analysis involves relatively simple and easily programmable calculations and its algorithms are clearly described by Lesage (1973).

The technique consists of selecting the two items which account for the highest correlation in a matrix as the "core" of a cluster. Each new member of the cluster is the one among the remaining items whose mean correlation with the items of the cluster is highest in absolute value. Once a cluster is closed, a new one is formed by repeating the process.

The conditions for closing a cluster and for stopping the process are the same. When a cluster is composed of only 2 items, the correlation between them must be at least equal to the correlation significant at the level which has been decided (e.g. $p < .05$). On the other hand, when a cluster is made up of more than 2 items, the criterion can be relaxed slightly. The condition then imposed is that, for an item to become the next member of the cluster, none of its correlations with the existing members of the cluster must account for less than half the variance which is accounted for by the correlation significant at the predetermined level. In other words, no correlation between the new member of the cluster and any of the existing members must be less than the square root of half the squared significant correlation.

There is a special condition which would lead to an item with the highest correlation mean being rejected from a cluster without the cluster being closed. Correlation means are calculated on absolute (signless) values. But, because correlations can be either positive or negative, the matrix of correlations for a cluster might well include both positive and negative signs. For a new item to be added to the cluster, it has to correlate consistently (i.e. always positively or always negatively) with those members of the cluster which correlate positively amongst themselves. Should it be found that it has some positive and some negative correlations with any items which are positively inter-correlated, then the item is discarded. The process then resumes on the same cluster with the next highest correlation mean.

To a certain extent, the contents of clusters already formed influence the contents of subsequent clusters. Once an item (i.e. construct) becomes a member of a cluster, it ceases to be eligible for subsequent clusters.

The process by which clusters are formed can be a source of errors. For example, there may be items which have high correlations with many others. Such items may become incorporated into an early cluster whereas other techniques would show them to belong to a smaller cluster. When this happens, the wrongly incorporated items in turn may artificially attract into the cluster items which would otherwise belong to other clusters.

Such potential sources of errors make it necessary to compare methods. When comparisons within and between samples are made, it is important that the dimensions which are compared are robust. The comparisons to be made within and between samples are described in the next sections of this chapter.

5.3 COMPARISONS WITHIN SAMPLES

Four samples - described in Chapter 6 - were selected for this study. Each sample rated different types of hazards. Within each sample, however, all respondents rated the same hazards as elements of their grid.

Comparing grids with common elements was thought to be the best way to test the hypothesis that hazard assessment patterns used by most people when considering the same hazards could be identified. Thus, consensus dimensions within these samples were sought. These consensus dimensions were then used for comparisons between samples in order to test the hypothesis that common hazard assessment patterns are used for most types of hazards (see 5.4 below).

In three of the four samples, respondents all rated the same elements on the same constructs. This made it possible to use statistical methods in order to verify whether consensus dimensions could be identified. How useful these statistical methods, described below, were was fully appreciated when an attempt was made to identify consensus dimensions in the third sample.

Respondents of the third sample all assessed the same hazards. In doing so, however, no two respondents used identical series of constructs. Constructs were elicited and grids were rated on an individual basis. Very few constructs were common to a majority of grids. As pointed out earlier it was expected to identify hazard assessment patterns underlying constructs. As constructs varied from one grid to another, it was difficult to ascertain whether dimensions extracted from different grids were similar.

The other type of dimension contents which could be compared were elements. There was, however, a greater degree of subjectivity involved in comparing dimensions on the basis of their elements than there was in comparing on the basis of constructs. When elements were compared, an inference had to be made to discern a characteristic which explained the observed polarisation of elements. A further inference had to be made to discern whether the characteristic was the same in two dimensions being compared. In dimensions compared on the basis of their constructs, these inferences were not needed since the characteristics, i.e. the constructs, were quite explicit.

Consensus dimensions were identified by Golant and Burton (1969) and by Champion (1977). However, it was argued in Chapter 4 that the reliability of these dimensions was not assessed. In other words, no comparisons were made to test whether consensus dimensions were faithful representations of dimensions identified in individual grids or tests. Furthermore, in both studies, only one method was used to derive consensus dimensions. Thus, there is a risk that these dimensions are statistical accidents; other methods for deriving consensus dimensions might have given different results.

In this research, three techniques were used to identify consensus dimensions. For the three samples in which grids were identical, these techniques made it possible to assess the reliability of the consensus dimensions which were identified.

5.3.1 Matrix of mean scores

The first method of deriving consensus dimensions which can be used for identical grids is based on a matrix of mean cell scores. This method was used by Golant and Burton (1969) and by Champion (1977).

A grid can be seen as a series of cells arranged in a two-dimensional matrix, one dimension being the constructs and the other the elements. The content of each cell is the rating of the corresponding element on each construct. When a given construct and a given element are used in more than one grid, a consensus score can be calculated by computing the mean rating of that cell in different grids. When, as in three of the four samples, all elements and all constructs are identical between grids a whole new grid made up of mean cell scores can be constituted.

Such grids can then be analysed in the same way as an individual grid. Matrices of mean cell scores (hereafter referred to as $\overline{X}$) were computed for three samples. Principal component analysis and cluster analysis were then performed on the three matrices.

5.3.2 Matrix of mean Z-scores

The same rating scale may sometimes be used differently by different people. Two people giving a rating of 4 to the same element on the same construct on a five-point scale may be doing so for quite different reasons. For instance, to a respondent who rated most elements as 2 a rating of 4 means an important difference, but to a respondent who rated most elements as 4, the same rating means no difference.

Therefore, means of raw scores, as in the matrix of mean scores, may be slightly misleading. It may prove valuable to construct a consensus grid based on the distance between an element and the construct mean, and the direction (positive or negative) of that distance, between grids. Distance and direction are both conveyed by Z-scores. The sign of the Z-score indicates whether the element is rated above or below the mean; and the absolute value of the score indicates how far from the mean the element is.

It was therefore decided to convert individual grids, in the three samples mentioned above, into matrices of construct Z-scores. The mean matrix of these grids was then calculated and processed as an individual grid (hereafter referred to as $\overline{Z}$).

5.3.3 Frequencies

Both methods described above make the same assumption that means (whether of raw scores or of Z-scores) are accurate representations of population scores. This may be true when variation about the mean is small. But when variation is large, means tend to obliterate individual differences.

As pointed out earlier, it was thought important to test whether consensus dimensions were accurate representations of dimensions found in individual grids. Therefore, it was decided to count the number of individual grids in which elements and constructs of consensus dimensions loaded significantly on corresponding individual dimensions. The number of individual grids in which a construct was retained as a member of a cluster corresponding to a consensus dimension was also counted. These various frequencies proved to be useful indicators of the robustness of consensus dimensions.

Other frequencies were also computed, but for different purposes. Firstly, whilst it is expected that underlying trends are identical across grids, these underlying trends could not be identified if no variance could be apportioned, in other words if ratings were totally unanimous. As pointed out in the discussion of Golant and Burton's results, unanimous ratings indicate that a test is not able to discriminate respondents or items. Therefore, the number of times that each rating was assigned to each element on each construct was computed. These frequencies were calculated in order to test whether individual differences in ratings could be found.

Secondly, the same frequencies were compared between samples. The object of the exercise was to test whether respondents in all samples used their rating scales in the same way.

5.3.4 Comparisons within the third sample

As pointed out earlier, respondents from the third sample used different constructs. Too few constructs were common to enough grids in the sample to make it worthwhile comparing grids reduced to common constructs.

One feature, however, was common to all grids in the sample: the elements. It was therefore possible to compare components between grids on the basis of the elements which loaded significantly on them. Of the three methods described above, only frequencies could be used to achieve such comparisons.

Some comparisons of constructs were possible. There were two possible ways of carrying out those comparisons. Firstly, one could perform a frequency count of the dimensions on which constructs occurring in at least a few grids loaded significantly. Secondly, labels may be affixed to the various dimensions of each individual grid and the labels compared between grids.

Both methods have disadvantages. Because the first method only involves a limited number of constructs, and because the number of common constructs varies from one pair of grids to another, the picture thus derived is likely to be incomplete and potentially misleading. As for the second method, the amount of subjectivity involved in it may cast serious doubt on its validity.

It is possible that the results of these two methods may not coincide with the results from other samples. Should that be the case, there should be indications as to whether differences are attributable to invalid methods or to altogether different patterns of hazard assessment. For instance, should both methods yield similar results, but these be different from other samples results, the likeliest explanation of the differences should be that there are different hazard assessment patterns. Should the results of only one of the comparison methods happen to coincide with those from other samples, the validity of the other mode might be considered dubious. Finally, should both methods yield results comparable to those from the other three samples, fewer doubts would remain about their validity.

5.3.5 Summary

The advantages of using statistical techniques for comparing identical grids were discussed earlier. There were specific reasons why three separate methods were retained. The matrix of mean scores was retained on the basis that it had been used in two other pieces of research. But it was felt that Z-scores were a more accurate description of relative ratings, so the method of mean Z-scores was used. The use of frequencies is a somewhat less robust method than the other two, although it does provide indications of individual variation. In addition, this method is not based on means and as such does not make the same assumptions as the other two methods.

It was therefore felt that the use of these three methods might be a precaution or a safeguarg against possible shortcomings in any one of them. In any method of extracting underlying trends, any dimension which is identified is almost invariably less sizeable or important (i.e. it explains less of the test variance) than the previous one. Therefore, the more dimensions there are, the more difficult it is to identify or extract the later ones. It was expected that the methods described above should be relatively unanimous about the first consensus dimensions but might show discrepancies about later dimensions.

On that basis, it was decided to incorporate into a consensus dimension those elements or constructs which were retained by at least two of the methods. Should the results of one method differ from those of the other two about an item (construct or element), then this can be taken to be as much an indication of a shortcoming in the method as of a difficulty in proving a relationship between item and dimension.

5.4 COMPARISONS BETWEEN SAMPLES

No statistical method could be identified which made comparisons between totally different grids possible. Numerical comparisons between samples were therefore of necessity limited. Those comparisons which could be made were similar in nature to those between individual grids in the third sample.

Firstly, consensus dimensions were compared on the basis of whether they encompassed constructs which were relatively comparable and, if so, how many such constructs were involved. Secondly, the labels affixed to the various consensus dimensions were compared. The aim of these comparisons was to find out whether certain labels were common to dimensions identified from different samples.

Of all the various levels of comparisons described above, those between samples undoubtedly involved the greatest degree of subjectivity. But by then, dimensions being compared were relatively robust and well defined. Thus, much as in clinical diagnosis, although uncertainty can never be removed entirely, it was hoped that indications, like symptoms, should converge to leave little doubt about the nature of the phenomenon of hazard assessment.

5.5 OPERATIONAL HYPOTHESES

A first formulation of the hypotheses of this research was stated in Chapter 4 following theoretical considerations. In view of the discussions in this chapter about measurement instruments and statistical techniques, it is now possible to propose more precise versions of the same hypotheses.

Firstly, the hypotheses about the nature of hazard assessment can be formulated to take the chosen measurement instrument and analysis technique into account. Thus:

- H1: When hazards are assessed as elements of a repertory grid, at least three underlying assessment patterns can be identified by principal component analysis.
- H2: The first and most important assessment pattern to be extracted concerns the controllability of hazards.
- H3: The second assessment pattern to be extracted concerns the severity of potential consequences of hazards.

H4: The third assessment pattern to be extracted concerns the likelihood of hazards leading to accidents

As pointed out in this chapter, it is expected that these emerging dimensions are not artefacts of the principal component analysis. Thus:

H5: These assessment patterns are robust and as such they can be identified both by cluster analysis and by principal component analysis.

Furthermore, it is postulated that the basic hazard assessment patterns are comparable between individuals considering the same hazards (within samples) and regardless of the types of hazards being considered. Thus:

H6: Most individual grids having identical elements yield identical hazard assessment patterns.

H7: Hazard assessment patterns extracted from consensus grids are similar across samples.

Finally, despite identical hazard assessment patterns, some individual differences are expected. Thus:

H8: Whilst emerging dimensions are identical across the grids of a sample, hazards can be construed differently by different respondents and hence ratings of elements on constructs are not unanimous.

The next chapter describes steps which were taken in order to test these hypotheses.

CHAPTER 6 : SAMPLES AND DATA COLLECTION

6.1 CHOICE OF SAMPLES

Having determined what measurement instrument should be used and how the data should be analysed, the next step was to choose samples of respondents who would provide the data. This choice is described in the present chapter.

Data had already been gathered from two samples by other researchers and permission was granted by them to use these data for this research. For methodological reasons, two more samples were included in this study. The way in which respondents in those two samples scored their grid is also described in this chapter.

6.1.1 First sample

It was mentioned in Chapter 4 that although Champion's (1977) results were interesting, some anbiguity remained in them. For instance, Champion relied solely on principal component analysis; thus, it is not known how robust are the components which were extracted. This analysis was only carried out on the grid of mean scores; some important individual differences may have been overlooked. Finally, Champion did not look into elements loading significantly on the four trends which he identified; it was argued earlier that the third and fourth trends were somewhat ambiguous.

The method described in Chapter 5 for the analysis of individual and consensus grids can overcome these problems. It was thought that, if Champion's data could be thoroughly re-analysed by that method, they would provide very useful information for the main investigation of this thesis.

It was decided to seek permission to re-process the grids scored by respondents from the beer kegging plant. This permission was kindly granted by Champion (1978). These respondents are hereafter referred to as GR respondents. Their elements were presented earlier in Table 4.5 and their constructs in Table 4.6 . For the sake of convenience these tables can be found in fold-out form in Appendices 1 and 2 respectively.

One of the methodological considerations which was raised in the discussion of Champion's project was that the constructs which participants used to rate hazards had been supplied by respondents in pilot studies. In other words, as was the case in Golant and Burton's (1969) research, respondents were using constructs which had been provided by someone else. This alteration of the usual grid procedure may have turned the grids into something more akin to semantic differential tests.

The main criticism of semantic differential tests is that some concepts (or elements) may be totally outside the range of convenience of their scales (or constructs). There are, however, indications that this criticism may not be fully applicable to Champion's grids. For instance, Champion's constructs appear to have greater face validity than Golant and Burton's scales. In other words, Champion's constructs seem to be more closely relevant to the elements of the grids. Furthermore, constructs in Champion's study were supplied by respondents (in pilot studies) rather than by the researcher. Such constructs are more likely than semantic differential scales to have been understood by those respondents who rated the grids. That only two respondents enquired from their interviewer about the meaning of constructs may be taken as an indication that constructs were easily understood by respondents who rated the grids.

However these considerations are insufficient to preclude the possibility that Champion's grids incorporated some of the same biases as those inherent in semantic differential tests. Comparisons with other grid elicitation techniques was necessary in order to test for biases.

6.1.2 Second sample

It was therefore decided to use another sample of respondents who would elicit constructs and, if possible, elements. An opportunity for doing so arose within the curriculum of the B.Sc. course in the Department of Safety and Hygiene. Second-year students from this course were to study the repertory grid technique. An experiment was devised to serve the twin purpose of being a practical example for the students and of collecting research data. Participants in this experiment are hereafter referred to as UG respondents.

A number of considerations were taken into account in the design of this experiment. Firstly, the aim of this research was to investigate the existence of general mechanisms of hazard assessment. In other words, assuming the pattern identified from Champion's data to be reliable, it had to be shown that the same pattern was also used in the assessment of other types of hazards.

Another aim of this thesis was to test the hypothesis that, within general assessment mechanisms, individual differences could be found. A number of considerations had to be taken into account in testing this hypothesis.

For instance, differences between grids could mean that different assessment patterns altogether are used by different respondents. One of the hypotheses of this research implies that individual differences between patterns should not be found. This hypothesis can be tested by comparing grids elicited individually. As mentioned later this was done in the third sample.

Individual differences postulated in this research are of another type. What is hypothesized is that within similar patterns ratings of hazards can be different between individuals. In order to identify such differences one has to have a number of respondents rate the same grid (same elements and same constructs). It can be argued that this was done in Champion's research. However, it has been suggested earlier in this chapter that differences found by Champion between individual grids may have been attributable, at least in part, to respondents' different understanding of constructs.

It was decided that, in order to ensure a relatively uniform understanding, constructs should be elicited through a group discussion. This group discussion is described in greater detail in the section on data collection in this chapter. It was thought that such a construct elicitation method would serve the purpose of making sure that all respondents would attach a similar meaning to a construct while enabling the researcher to have an adequate understanding of that meaning. Furthermore, as pointed out in Chapter 5, it was thought that an exchange of ideas in a group discussion would ensure that no salient or relevant construct would be omitted through oversight.

Although closer than a standard semantic differential test, for example, to Kelly's repertory grid technique, the method described above probably did, to a certain extent, impose a fixed framework upon respondents. A way had to be found of testing the hypothesis that hazard assessment patterns identified in the first two samples were not artefacts of fixed grids and could be identified in individually elicited grids.

6.1.3 Third sample

Both samples described so far had three characteristics in common. Firstly, within each sample, all respondents used a fixed grid, i.e. the same elements and constructs were used by all respondents within a sample. It can be argued that a fixed grid created constraints which respondents had to cope with and that, consequently, the results were influenced by those constraints.

Secondly, in both experiments construct elicitation and grid scoring were unrelated. Undergraduate students, as will be explained later, scored their grids a few weeks after the construct elicitation session. Respondents in Champion's study were using constructs elicited by others.

Thirdly, respondents in both groups had a relatively strong interest in health and safety matters. Participants at the beer kegging plant were mostly members of the safety committee; the second sample was made up of students in the second year of a three year course in occupational health and safety. It is therefore possible that, as a result, both samples were biased and that respondents' perceptions of hazard differed both from each other and from those of the population at large.

Therefore, it was thought necessary to include another sample of respondents without any active involvement in occupational safety and health matters. Such a group should be subjected to the standard grid technique and score their grid immediately after construct elicitation.

Respondents in Green and Brown's (1976 b, 1977 b) EI experiment met these three conditions. In addition, the data provided by these respondents had not been analysed in any great detail. Green (1978) kindly agreed to the data being analysed and to the analysis of results being incorporated into this thesis. Participants from that sample are hereafter referred to as SA respondents. Their elements were listed in Table 4.3 and can be found in fold-out form in Appendix 5.

6.1.4 Fourth sample

Additional data were obtained as part of another study conducted by Champion and myself. The research was concerned with an investigation into the psychology of pot-holing. Part of this research consisted of a repertory grid where the elements consisted of pot-holing hazards and dangers.

In a group discussion three experienced cavers were asked to mention what they thought were the hazards of pot-holing. Nine elements were thus elicited. The usual construct elicitation technique of comparing elements in triads yielded six constructs.

It can be argued that only three respondents rating a very small grid is not likely to provide much useful information. But this grid combined two methodological features which have been discussed previously. Firstly, elements and constructs were elicited in a group session with the three respondents. Secondly, the grids were scored immediately after construct elicitation. Of the other three samples, only in the SA sample were the grids scored immediately after construct elicitation; but the technique was conducted on an individual basis. The other two samples used grids elicited in group sessions; but some time elapsed between construct elicitation and grid scoring.

Because of the sizes of the sample and of the grid, it may prove difficult to adequately compare assessment of pot-holing hazards with assessment of other types of hazards. But this fourth sample can shed some light on differences which may exist between samples and which can be attributable to differences in grid techniques.

6.2 DATA COLLECTION

The way in which grids were elicited, assembled and scored by respondents in the first (Champion's) and the third (Green's) samples has been described in Chapter 4 and earlier in this chapter. But, so far, only sketchy details have been provided about the second and fourth samples and the grid used by these respondents. This section is a more complete description of the way in which data from the second and fourth samples were gathered.

6.2.1 Home hazards as elements

The first decision which had to be made in this experiment was the choice of stimulus material. In other words, hazards had to be chosen as elements of the grid.

One of the hypotheses of this thesis states that some hazard assessment patterns are applicable to most types of hazards. Some types of hazards had already been used in other samples. For instance Champion (1977) had listed industrial hazards as elements; Green and Brown (1976b, 1977b) had used natural and large-scale hazards in their EI experiment. In order to perform a more elaborate test of the hypothesis that the basic hazard assessment patterns are applied to all types of hazards, it was decided to use as elements in this sample a different type of hazards from those used in the other two samples.

Stimulus material related to home hazards was readily available from another experiment. De-briefing interviews were conducted as part of the diary techniaue experiments described earlier.

Part of these interviews consisted of a game-like exercise; respondents were asked to identify all the hazards they could find in a drawing of a scene in a kitchen. The picture was devised by the Royal Society for the Prevention of Accidents (RoSPA) for use in a "Spot the Hazard" competition. Permission had been sought from RoSPA to use the picture for research purposes; Mrs E. Maclean, director of the Home Safety section, granted that permission on behalf of RoSPA.

It was decided to use that picture (hereafter referred to as HS-CP9) for the element elicitation part of the grid technique. During the lecture period allocated to this exercise, all undergraduate respondents were given a copy of HS-CP9. They were asked to mark on the picture the hazards they thought were present in the scene. They were given ten minutes to complete that part of the exercise. Then, the pictures were gathered. A group discussion followed on the hazards which they had idetified; following the discussion, a list of 38 elements was drawn up. These can be found in Table 6.1 and in Appendix 3. This list in itself is interesting since RoSPA, in their list of correct answers to HS-CP9, mention only 17 items.

6.2.2 Construct elicitation

Construct elicitation was undertaken as soon as the list of elements was complete. Respondents were asked to consider the first three elements on the list and to suggest ways in which two were similar and different from a third. All constructs elicited from this first triad were noted before moving on to the next triad.

TABLE 6.1

LIST OF HAZARDS USED AS ELEMENTS FOR THE GRID

RATED BY UG RESPONDENTS

1. Too many things going on at once.
2. Ball in the way.
3. Rucked up mat.
4. Soap suds on the floor.
5. Bucket in the way.
6. Knife sticking out of the drawer.
7. Tin left on the floor.
8. Jagged edge on tin lid.
9. Flip flop shoes.
10. Matches on cooker.
11. Kettle lead in water.
12. Faulty wire on mixer.
13. Confusion over on and off position of switches.
14. Fat boiling over.
15. Mixer overhanging the working surface.
16. Cooker switches hard to reach.
17. Two (2) cooker rings on.
18. Kettle (electric) left on cooker ring.
19. The fact that the drawer is open.
20. Lady not looking at what she is doing.
21. The way the lady is holding the vegetable for cutting.
22. Improperly stacked bottles in cupboard
23. Door of cupboard is open.
24. Position of the bin.
25. The fact that the bin is overfilled.
26. Broken bottle sticking out of the bin.
27. The man's vision being obscured by the box.
28. Box with fragile contents, improperly sealed.
29. The way the man is holding the box.
30. Switch on mixer left "on".
31. Untidy work surface (loose vegetables, etc.)
32. Lady not wearing an apron.
33. Tin opener left lying about.
34. Edge of botton cupboard.
35. Edge of top cupboard.
36. Stacking of cups and saucers.
37. Position of cooker as regards opening of oven door.
38. Position of the chip pan handle.

Each element was used only once. Thus, the second triad included elements numbers 4, 5 and 6. This way 38 elements represented 12 triads, the last two elements being left out. By the time the twelfth triad was reached, because many individuals had volunteered constructs on each preceding triad practically all the last constructs were repetitions of already mentioned ones. Thus, there was no point in continuing with rearranged triads. Construct elicitation being over, a discussion took place about the exercise.

Some criticisms were made by respondents during this last discussion. They were mostly directed at the way in which constructs were elicited. The main argument was that elements were listed by categories. For instance, a fire hazard was mentioned by someone. This would generate a series of fire hazards, mentioned by other respondents. When triads were discussed,at least one would be made up entirely of fire hazards. As a consequence, according to respondents, triads made up of only one type of hazard limited the range of elicited constructs. In other words, they argued that a greater variety of elements in a triad would have led to a greater variety of constructs. The rest of the criticisms were directed at the tediousness of eliciting a grid of 32 constructs from 38 elements.

The whole of the construct elicitation discussion and most of the subsequent discussion were recorded and the recording was later studied in detail. The purpose of that scrutiny was twofold. Firstly, it served as a check that all constructs which had been mentioned appeared in the final list. Secondly, attention was paid to the way in which constructs were mentioned.

The aim of the experiment was for respondents to suggest constructs for the grid. But in so doing, they could have embarked upon a discussion which could have led to the discarding of constructs on which there was not unanimous agreement. It was not the purpose of a group construct elicitation session to retain only consensus constructs or to arrive at a consensus rating of hazards. The discussion was therefore recorded and subsequently re-analysed to check whether these pitfalls were avoided. The result of the scrutiny satisfied the experimenter that, during construct elicitation, respondents were in fact mentioning criteria or scales for assessing hazards rather than collectively assessing the hazards.

6.2.3 Grid preparation

Before the grid was presented to respondents for scoring, some preparation work was done. Firstly, the technical constructs were removed. As it stood, the grid was nearly four times the size of Champion's grid; and yet, the latter had taken respondents an hour to score. It was therefore decided to try and reduce the size of the grid without imposing the experimenter's framework on the grid and without wasting opportunities of collecting valuable information. That is why technical constructs such as "electrical hazard - not electrical hazard" were not included in the final grid. This reduction of the number of constructs from 32 to 27 reduced the grid to three times the size of Champion's grid.

The second type of preparation was the formulation of as many constructs as possible for use with a five-point scale ranging from "Strongly disagree" to "Strongly agree". Sixteen constructs were thus arranged. The aim of this conversion was to facilitate scoring by removing the hurdle of scale interpretation for each construct. As for the remaining eleven constructs, they were also arranged in the form of a statement with five choices of reply. However, these choices differed from the first sixteen constructs. The final list of constructs and their rating scales can be found in Table 6.2 and in Appendix 4.

Anchoring all five points of the rating scales was done for comparison purposes. In Champion's GR experiment and in the pot-holing study scales were anchored only at both ends. In Green and Brown's EI experiment the mid-point was also labelled. There is a possibility that respondents tended to use mostly those points of a scale which were labelled. It was decided to label each point of each scale in order to see if respondents from the second sample used rating scales differently from other respondents.

The final grid was presented as a pencil-and-paper questionnaire. Each page of the questionnaire consisted of the list of elements and was headed by one of the constructs. The order of pages was the same in all questionnaires. In parallel with the list of elements, there were five columns corresponding to the five points of the rating scales. Rating of an element was done by placing a tick or a mark in the appropriate column next to that element. Then, to complete the questionnaire, a new copy of HS-CP9 was attached. This was done in order to help respondents remember exactly the nature of the elements they were to rate.

TABLE 6.2

LIST OF CONSTRUCTS AND THEIR SCALE IN THE UG GRID

CONSTRUCT NUMBER	DESCRIPTION	SCALE
1	This is the consequence of another hazard	Strongly agree / Strongly disagree
2	This is a hazard	Strongly agree / Strongly disagree
3	This is due to inadequate maintenance	Strongly agree / Strongly disagree
4	This hazard is only temporary	Strongly agree / Strongly disagree
5	This is the result of untidiness	Strongly agree / Strongly disagree
6	In order to call this a hazard, one has to make some assumptions	Strongly agree / Strongly disagree
7	This is caused by a lack of knowledge	Strongly agree / Strongly disagree
8	This is sufficient on its own to cause an injury	Strongly agree / Strongly disagree
9	This is the result of hurrying	Strongly agree / Strongly disagree
10	It takes a specialist to put this right	Strongly agree / Strongly disagree
11	This arises from something that is in use	Strongly agree / Strongly disagree
12	This is the result of poor planning	Strongly agree / Strongly disagree
13	This is a multiple hazard	Strongly agree / Strongly disagree
14	This arises from something not having been done	Strongly agree / Strongly disagree
15	This is due to bad design	Strongly agree / Strongly disagree
16	This is the result of someone's action(s)	Strongly agree / Strongly disagree

TABLE 6.2 (Continued)

No.	DESCRIPTION	SCALE
17	How hard/easy is it to see this?	Very easy / Very hard
18	How easy/difficult is it to remove this?	Very difficult / Very easy
19	How easy/difficult is it to avoid this?	Very difficult / Very easy
20	How easy/difficult is it to put this right?	Very difficult / Very easy
21	How serious/trivial an injury could this cause?	Very trivial / Very serious
22	How likely/unlikely is this to cause an accident?	Very unlikely / Very likely
23	How direct/indirect a cause of injury can this be?	Very indirect / Very direct
24	How large/small a hazard is this?	Very small / Very large
25	How dangerous is this?	Very dangerous / Not dangerous at all
26	One comes across this...	Very frequently / Very infrequently
27	How likely/unlikely is this to cause death?	Very likely / Very unlikely

6.2.4 Grid scoring

A second lecture period had been scheduled for the exercise. This second hour was devoted to grid scoring. The first five minutes were taken to explain how to score the grid. By the end of the hour, those respondents who worked fastest had completed half the task. During that time, questions of clarification about elements were answered. But questions about the meaning of constructs were parried on the grounds that what was asked for was the respondent's interpretation, not the experimenter's. Most of these questions related to construct 13.

Respondents took the questionnaire away with them at the end of the hour and completed it in their own time. The first questionnaire to be handed back was returned two days later; the last one was recovered six weeks later. Whenever a grid was handed back, it was checked for missing values and when these were found, the respondent was asked to fill in the blanks. Missing values varied in numbers from zero to eleven. In view of the low numbers, no pattern of errors was discernable.

There were, however, missing constructs. Two respondents omitted construct 13 completely on the grounds that they did not understand it. For the same reason, one respondent omitted constructs 11 and 21 completely. Construct 13 appears to have been an awkward one for respondents to use. But all these problems will be looked into when the results of that experiment are discussed.

6.2.5 Caving hazards

As pointed out earlier, the three respondents of the fourth sample were experienced speleologists. They were interviewed a few hours after having been on an expedition. At first, the group discussion centered around their expedition. Then, they were asked to name the hazards of caving. The hazards they named, listed in Table 6.3 and in Appendix 6, were then used as elements.

Elements were arranged in triads in the usual way. In other words, the first triad included elements 1, 2 and 3, the second triad elements 2, 3 and 4, and so on. The last triad included elements 9, 1 and 2. Respondents were allowed to name as many constructs as they wished to for each triad. A total of six constructs were thus elicited, all constructs from later triads being repetitions of already mentionned constructs. The final list of constructs can be found in Table 6.4 and in Appendix 7.

The constructs were presented as bi-polar five-point rating scales. Respondents were told that left-hand-side labels represented a rating of 1 whilst right-hand-side labels corresponded to ratings of 5.

As the grids were small, they were rated there and then. The longest it took to assign the 54 ratings was 10 minutes.

TABLE 6.3

LIST OF HAZARDS USED AS ELEMENTS BY
PH RESPONDENTS

1. Failure to take space lighting on long trip.
2. Insufficient care taken when climbing dislogded rocks.
3. Dislodged miners' deads in a mine.
4. Digging in an unstable choke without shoring.
5. Bad air in airbell between sumps.
6. Inexperience of party especially leader.
7. Rope abrasion on SRT rope.
8. Loose rocks in boulder choke.
9. Poorly supervised novices.

TABLE 6.4

LIST OF CONSTRUCTS USED BY

PH RESPONDENTS

1.	Controllable	-	Uncontrollable
2.	Lack of thought	-	Bad luck
3.	Due to my party	-	Due to other party
4.	Technical factor	-	Physical factor
5.	Poor planning	-	Poor moving
6.	Likely to kill	-	Unlikely to kill

CHAPTER 7 : INFLUENCES UPON RATING FREQUENCIES

Five-point rating scales were used with all four groups of respondents. There were nevertheless differences in the presentation of the scales. For instance, scales labelled only at the ends were used with the GR and PH samples. SA respondents used scales labelled at the mid-point as well as at both ends. All points of the scales used by UG respondents were labelled.

Many factors can influence rating. For instance, it is pointed out later in this chapter that the greater complexity of a person's thought processes, the more points of the rating scale are used. Another example is quoted by Edwards (1957); he argues that labelling each point of a rating scale helps respondents to understand their meaning and thus ensures more evenly distributed rati frequencies.

In turn, rating frequencies can have an important influence upon correlations and thus on the extraction of underlying trends. The variance of a measurement (e.g. a construct) is calculated from rating frequencies (Guilford and Fruchter, 1973). Correlations are measures of common variance. Data reduction methods identify underlying trends by apportionment of all common variance measured by correlations among a set of data.

Thus, when hazard assessment patterns are examined, it is important to look into the rating frequencies which yielded them. Rating frequencies and variables possibly affecting them are therefore considered in this chapter before proceeding, in subsequent chapters, to a discussion of hazard assessment patterns.

7.1 FIRST SAMPLE (GR)

When GR grids were processed, it was found that Ingrid 72 rejected a total of six constructs. One individual grid accounted for three of these rejections and another three grids each accounted for one rejected construct. In Chapter 5, it was pointed out that constructs are rejected when they have a null variance. Null variance occurs when a respondent assigns an identical rating to all elements on a construct. This construct is then rejected as null variance would make it impossible to calculate correlations with other constructs and might introduce bias into extracted dimensions.

Rating frequencies were analysed in a search for clues as to why constructs were rejected. Table 7.1 is a summary of that analysis. For each construct Table 7.1 lists the number of times that each rating was used for all elements in all grids, and means and standard deviations of the constructs. These statistics are also listed for total frequencies on 17 constructs.

Calculations for constructs 4, 7, 9 and 12, and for total, were done twice. The first calculation included all ratings. There were grounds, however, for believing that ratings from rejected constructs might introduce bias into the calculations. It could be, for example, that for certain respondents some constructs could not be used to sort out the elements; including these constructs with those involving a certain level of discrimination seemed somewhat awkward. Therefore, calculations were performed a second time by excluding ratings from

TABLE 7.1

SUMMARY OF RATING FREQUENCIES, MEAN AND STANDARD DEVIATION FOR CONSTRUCTS IN GR GRIDS

Construct	Rating					Mean	S.D.
	1	2	3	4	5		
1	65	19	46	33	65	3.06	1.58
2	45	45	41	18	79	3.18	1.56
3	100	25	32	15	56	2.57	1.65
4	74	9	17	15	113	3.37	1.80
*	74	9	17	15	94	3.22	1.81
5	125	30	40	10	23	2.02	1.34
6	138	22	22	14	32	2.04	1.49
7	116	34	42	19	17	2.07	1.30
*	97	34	42	19	17	2.16	1.32
8	124	36	37	13	18	1.97	1.28
9	11	19	60	39	99	3.69	1.21
*	11	19	60	39	80	3.57	1.22
10	54	36	71	46	21	2.75	1.27
11	41	44	64	25	54	3.03	1.40
12	176	7	29	5	11	1.54	1.06
*	119	7	29	5	11	1.72	1.22
13	70	25	38	11	84	3.06	1.69
14	43	12	22	21	130	3.80	1.59
15	65	14	61	35	53	2.99	1.51
16	111	13	40	11	53	2.48	1.65
17	77	28	44	14	65	2.83	1.63
TOTAL	1435	418	706	344	973	2.74	1.62
*	1359	418	706	344	935	2.75	1.61

* Calculated by excluding frequencies from rejected constructs.

rejected constructs.

7.1.1 Uneven ratings

Total frequencies in Table 7.1 show that the number of 1's ticked by respondents was nearly one and a half times the number of 5's (the second highest frequency) and approximately four times the number of 4's (the lowest frequency). Thus, ratings appear to have been assigned unevenly.

The null hypothesis that all scale points were ticked evenly was tested by a chi-square test. For corrected total frequencies, the chi-square value was 906.55 ($df = 4$, $p \ll .001$; see Table 7.2). The test result indicates that rating frequencies are significantly different from a random distribution. This difference is even greater when uncorrected frequencies are used.

It is interesting to note that the three most frequently ticked ratings were 1, 3 and 5, i.e. the points on construct scales which were labelled and the central point. This topic is re-examined later when the effects of labelling on rating frequencies are discussed. However, at least two other explanations can be suggested which might account for these differences in the frequencies.

Firstly, there may have been a prevalence of constructs on which the "natural" answer was a rating of 1. In other words, there may have been a number of constructs which were unbalanced towards one end. This may have been the case since 1's accounted for the highest frequency in 9 of the 17 constructs. There are, however, indications that the main factor may have been labelling only the end

TABLE 7.2

CHI-SQUARE TEST TO COMPARE TOTAL OBSERVED GR RATING FREQUENCIES AND THEORETICALLY EVEN FREQUENCIES

Rating	Observed frequencies	Expected frequencies
1	1359	752.4
2	418	752.4
3	706	752.4
4	344	752.4
5	935	752.4

$X^2 = 906.55$

$df = 4$

$p \ll .001$

points. For instance, 1's and 5's jointly accounted for the highest frequency on construct 1; for a further five constructs, 5's were used most often. For the remaining two constructs, the highest frequency was for 3's.

A second possible explanation for the discrepancy between frequencies is outlined by Edwards (1957). He states that, when a test item offers many choices of answers and when some ambiguity may exist between choices, there is a small but noticeable tendency for respondents to select the first choice.

A Mann-Whitney U-test was used to test for the presence of such a tendency. The assumption was made that, if the tendency did have an effect, highest frequencies would be greater for 1's than for any other rating. Results of the test are shown in Table 7.3; they only just fail to be significant at the .05 level. It therefore appears that if there was a tendency for respondents to selec the first choice of answer (i.e. a rating of 1) it did not significantly influence rating.

7.1.2 Influence of respondent's job

As pointed out earlier, Ingrid 72 rejected a total of six constructs. Three constructs were rejected only once, but construct 12 ('Preventable - unpreventable') was rejected from further analyses in three of the twelve grids. In each of these three grids, all elements were rated as 1 on construct 12. An attempt was made to discover why these unanimous ratings occurred.

TABLE 7.3

MANN-WHITNEY U-TEST ON HIGHEST GR FREQUENCIES

Rank	Construct *	1's	Others **
1	11		64(3)
2	15	65	
3	10		71(3)
4	17	77	
5	2		79(5)
6	9		80(5)
7	13		84(5)
8	4		94(5)
9	7	97	
10	3	100	
11	16	111	
12	12	119	
13	8	124	
14	5	125	
15	14		130(5)
16	6	138	
	Sum of ranks =	91	45
	U statistic =	17	46
	$p < .10$		

* : Construct 1 was not included as the number of 1's and the number of 5's were identical.

** : The figures in brackets are the ratings whose frequencies are listed.

Table 7.1 shows that, out of 228 ratings (12 respondents each scoring 19 elements) on construct 12, 176 were 1's. This is the highest frequency of any rating on any construct. Removing the ratings from the three rejected constructs leaves a frequency of 119. Despite being obtained from only 9 grids instead of 12, 119 still represents the fifth highest frequency for any rating on any construct. Therefore, there appears to be a marked tendency among the remaining nine respondents also to rate most elements as preventable.

Details of respondents' jobs were examined for further possible explanations of the three rejections of construct 12. The grids which showed no variation on this construct were rated by two security officers and a kitchen chef. This suggested the hypothesis that people not directly involved in the production process were more likely than people involved in the process to perceive hazards as preventable.

This hypothesis was tested by sorting respondents into those involved (2 fitters, 1 machine shop worker and 2 beer processing workers) and those not directly involved (4 security officers, 1 administrator, 1 storeman and 1 chef) in the production process. A Mann-Whitney U-test was applied on the number of 1 ratings provided by each on construct 12. The hypothesis that there was no difference between the groups was rejected at the .002 level of significance (cf. Table 7.4).

TABLE 7.4

MANN-WHITNEY U-TEST ON FREQUENCIES OF 1 RATINGS

ON CONSTRUCT 12 OF GR GRIDS

Rank	Grid	Involved	Not involved
1	GR09		8
2.5	GR02		11
2.5	GR07	11	
4	GR06	12	
5.5	GR10	14	
5.5	GR12	14	
7	GR03	15	
8	GR08		16
9	GR04		18
11	GR01		19
11	GR05		19
11	GR11		19
	SUM OF RANKS	24.5	53.5
	U STATISTIC	25.5	9.5
	$p < .002$		

What appears to be a general tendency to rate industrial hazards as preventable and the fact that this tendency appears to be more prominent in people not directly involved in the production process is worth further investigation, but is outside the scope of this thesis. These results, however, contribute to an explanation why construct 12 was rejected three times from further analysis by Ingrid 72. No evidence could be found for a respondent's job influencing the rejection of the other three constructs (number 4: 'Moving danger-stationary danger'; number 7: 'Very likely to cause an accident - very unlikely to cause an accident'; and number 9: 'Only a trivial injury - permanent disability').

7.1.3 High construct variation and the two-point scale

When rejections were investigated, it was found that three constructs were rejected from the grid scored by respondent GR 11. The way in which this respondent used the scales was examined for explanations of this triple rejection. The first clue was found in the mean variation per construct.

Variation on a construct is the sum of squared differences between the construct mean and each element rating. As mentioned in Chapter 5, Ingrid 72 prints out the variation for each construct in a grid. Constructs are rejected when they show a variation equal to zero. But a mean variation can be calculated for the remaining constructs.

The mean variation for those constructs which were not rejected was calculated for each grid. The highest mean was found to be that for grid GR11. The mean was found to be 25% higher than the next highest mean and more than three times the value of the lowest mean. Slater (1972) points out that high variations are often attributable to a respondent using mostly the extremes of rating scales. The hypothesis that this happened in grid GR11 was tested in two ways.

Firstly, the rating frequencies in grid GR11 were compared with the total corrected frequencies for the rest of the sample. The figures used in the calculation of the chi-square are shown in Table 7.5. The figures show that respondent GR11 used a significantly ($p < .001$) greater proportion of extreme ratings (1's and 5's) than did other respondents in the same sample. The proportion of 2's, 3's and 4's was therefore lower in grid GR11 than in other grids.

The second test of the use of extreme ratings was done by comparing rating frequencies in grids GR11 and GR12. Grid GR12 showed the second highest mean variation per construct. Frequencies were calculated only for the 14 constructs which showed variation about their mean in both grids. The results of the comparison are shown in Table 7.6. It can be seen that there were significantly ($p < .001$) more extreme ratings and fewer 3's in grid GR11 than in grid GR12.

The comparisons in Tables 7.5 and 7.6 are both strong indications that respondent GR11 used constructs as two-point scales.

TABLE 7.5

COMPARISON OF RATING FREQUENCIES BETWEEN

GR11 AND THE REST OF THE GR SAMPLE

RATING	GR11 Observed	GR11 Expected	OTHERS Observed	OTHERS Expected	TOTAL
1	129	96.1	1230	1259.3	1359
2	13	29.6	405	387.3	418
3	11	49.9	695	654.2	706
4	5	24.3	339	318.8	344
5	108	66.1	827	866.4	935
TOTAL	266		3486		3762

$X^2 = 99.89$

$df = 4$

$p < .001$

TABLE 7.6

COMPARISON OF RATING FREQUENCIES BETWEEN GR11 AND GR12

RATING	GR11 Observed	GR11 Expected	GR12 Observed	GR12 Expected	TOTAL
1	129	120.5	112	120.5	241
2	13	9.5	6	9.5	19
3	11	29.5	48	29.5	59
4	5	5	5	5	10
5	108	101.5	95	101.5	203
TOTAL	266		266		532

$X^2 = 27.81$

$df = 4$

$p < .001$

Thus, it seems plausible that constructs on which all elements would otherwise have received ratings of 1 or 2, for example, received unanimous ratings of 1 from respondent GR11 because he tended to use only the end points of rating scales.

As pointed out earlier, the proportion of extreme ratings was comparatively large in GR grids. It is possible that extreme ratings were preferred because they corresponded to the only two labelled points of the scales. This postulate can be corroborated later by considering how scales were used by other samples.

7.1.4 Other explanations

Although such an explanation seems less likely than one based on labelled points, it is possible that constructs which elicited unanimous ratings were not understood by respondents. If this were true, ambiguous or unclear constructs would be expected to yield ratings of 3, even if the mid-point is not labelled (Edwards, 1957). What was found in GR grids was a much higher frequency of 1's and 5's than of 3's.

Another explanation can be suggested to account for constructs yielding unanimous ratings. It is possible that constructs which yielded some identical ratings for the first few elements led respondents to adopt a response set; such a response set might in turn have led respondents either to overlook some of the words in certain elements or to read the descriptions less thoroughly.

Some indications of a response set are found in construct 4 ('Moving danger - stationary danger') of grid GR11. Respondent GR11 assigned a rating of 5 to all elements even though for 3 elements he was the only respondent to have assigned such a rating; and for one of those 3 elements all other respondents had assigned a rating of 1. Although the influence of a mental set on grid ratings cannot be tested further with the present data, such an influence cannot be rejected as a possible explanation of unanimous ratings.

Further explanations as to why 3 constructs were rejected from grid GR11 were sought in the nature of the rejected constructs. An analysis of the ratings in that grid revealed that the respondent rated all elements as stationary, very likely to cause an accident and preventable. If such ratings have something in common, the underlying characteristic is far from obvious.

But of all GR respondents, respondent GR11 was probably the least directly involved in the production process; he was the kitchen chef. It is therefore possible that his direct experience or knowledge, and hence his understanding of the elements was less precise than other respondents'. It is for example surprising that a fautly fork-lift truck in operation should be rated as stationary. Whether such unusual ratings are attributable to response set or to inaccurate understanding is difficult to say.

So far, explanations have been suggested for construct 12 being rejected in 3 grids and for respondent GR11 assigning identical ratings to all elements on a construct. One rejection remains to be explained: that of construct 9 ('Only a trivial injury - permanent disability') in grid GR10. No specific clues were found as to what possible explanation for that rejection was the most plausible. Respondent GR10 assigned ratings of 1 slightly more often in general than did other respondents but this tendency was found not to be significant (chi-square = 4.31, df = 4, $.30 < p < .50$). In any case, the rejected construct yielded constant ratings of 5. Analyses of rating frequencies yielded no more clues about alternative explanations.

7.2 SECOND SAMPLE (UG)

Whereas constructs had been rejected by Ingrid 72 in the GR sample, a somewhat different problem was found in UG grids. One construct was actually missing from two of the UG grids and two constructs were missing from another grid. Two respondents mentioned not having understood construct 13 ('This is a multiple hazard'). Another respondent mentioned not having understood constructs 11 ("This arises from something that is in use") and 21 ("How serious/trivial an injury could this cause?"). Accordingly, these three constructs were not rated by these three respondents.

Constructs 11 and 13 are among the more obscurely worded constructs of the UG grids and as a result, there may have been an ambiguity problem, perhaps additionally complicated by a language barrier. For the three respondents who omitted constructs, English was not a first language.

7.2.1 Uneven ratings

Table 7.7 displays the ratings on the various constructs in UG grids. The mean and standard deviation of each construct are also listed. The totals show that 2's and 4's were the most frequently used ratings.

Table 7.8 shows the results of the test of evenness of rating distribution. There were significantly ($p \ll .001$) more 2's and 4's and fewer 1's, 3's and 5's than if ratings had been assigned evenly. The problem, encountered in GR grids, of artificially high variance is therefore less likely to have occurred in UG grids.

There can be the alternative problem that low variance could lead to inadequate extraction of underlying dimension. In general the standard deviation of UG constructs was smaller than that of GR constructs. However, it is argued in Chapter 9 that low variance did not represent much of a problem. As for the predominance of 2's and 4's, an explanation is suggested in section 7.5.2 where the effects of labelling are discussed.

7.2.2 Response pattern in constructs

The order of presentation of the constructs in GR grids was randomized. There was therefore very little chance of consensus dimensions being affected by a response pattern. However, because the order of constructs was identical in all UG grids, it was important to ensure that there was no response pattern which would affect the extraction of consensus dimensions for this sample.

TABLE 7.7

SUMMARY OF RATING FREQUENCIES, MEAN AND STANDARD DEVIATION OF CONSTRUCTS IN UG GRIDS

CONS-TRUCT	FREQUENCY OF RATINGS 1	2	3	4	5	MEAN	S.D.
1	53	138	93	107	27	2.80	1.14
2	108	160	93	45	12	2.27	1.05
3	60	113	69	109	67	3.02	1.32
4	38	172	75	76	57	2.86	1.18
5	96	140	70	41	71	2.64	1.38
6	50	171	72	82	43	2.75	1.20
7	32	83	83	128	92	3.39	1.24
8	98	102	88	100	30	2.18	1.36
9	64	143	75	66	70	2.84	1.33
10	30	24	20	97	247	4.21	1.21
11	49	120	65	77	69	2.99	1.33
12	52	135	96	82	53	2.88	1.23
13	27	141	48	108	18	2.71	1.12
14	112	165	63	57	21	2.31	1.15
15	41	58	60	150	109	3.54	1.28
16	134	108	75	79	22	2.39	1.25
17	100	155	83	57	23	2.40	1.15
18	34	42	58	112	172	3.83	1.29
19	25	58	72	162	101	3.61	1.16
20	28	48	51	144	147	3.80	1.22
21	24	41	87	142	86	3.59	1.14
22	36	42	111	156	73	3.45	1.15
23	40	112	87	115	64	3.12	1.23
24	43	79	134	112	50	3.11	1.16
25	53	136	134	61	34	2.73	1.11
26	76	132	125	63	22	2.58	1.11
27	32	32	76	147	131	3.75	1.19
TOTAL	1535	2850	2163	2675	1911	3.05	1.32

TABLE 7.8

TEST FOR DIFFERENCES BETWEEN OBSERVED FREQUENCIES AND FREQUENCIES FOR A THEORETICALLY EVEN USE OF THE UG RATING SCALE

RATING	FREQUENCIES	
	Observed	Expected
1	1535	2226.8
2	2850	2226.8
3	2163	2226.8
4	2675	2226.8
5	1911	2226.8
TOTAL	11134	

$X^2 = 526.16$

$df = 4$

$p \ll .001$

The split-half technique was used to test whether the rating scales were used in the same way for all constructs. When the rating frequencies of the first thirteen constructs were compared to those of constructs 15 to 27, a significant difference was found ($X^2 = 231.68$, df = 4, $p < .001$). This finding was checked using Edwards' (1957) technique. The rating frequencies of the first six and of the last six constructs were compared. Again, a significant difference was found ($X^2 = 188.53$, df = 4, $p < .001$).

These tests revealed greater frequencies of 1's and 2's in the first constructs and of 4's and 5's in the last constructs. However, when Mann-Whitney U-tests were calculated to compare construct variances both between the first and the last half of the grid and between the first and the last six constructs, no significant difference was found. Thus it seems that a likely explanation for the differences found is that constructs suggesting ratings of 1 or 2 prevailed at the beginning of the grid and that constructs where a popular choice of answer was 4 or 5 were to be found towards the end of the grid.

Constructs at each end of the grid may not have been typical of the whold grid for a number of reasons. For instance, it is possible that respondents became familiar with the rating process only after rating a few constructs. It is also possible that because of the length of the grid a fatigue factor might be reflected by the rating of constructs either in the middle or at the end of the grid. A variant upon this possibility is that a positive goal gradient effect overrode fatigue for the last few constructs of the grid.

Some statistical checks were performed to test for the presence of such effects; these checks were interpreted as two-tail. Firstly the rating frequencies of the six constructs at each end of the grid were compared with those of the middle fifteen constructs. Each end was found to be significantly different from the middle; when the first six constructs were compared, the chi-square was 242.36 ($df = 4$, $p < .001$) and when the last six constructs were compared, the chi-square was 131.93 ($df = 4$, $p < .001$). It was pointed out earlier that there was a distinct prevalence of constructs eliciting ratings of 2 at the beginning of the grid and ratings of 4 towards the end of the grid; the tests mentioned above indicated that there was an even mixture of both types of constructs in the middle of the grid.

Secondly, the variances of the six constructs at each end were also compared with those of constructs 7 to 21. Neither end of the grid could be shown to have variances significantly different from the variances in the middle of the grid. These results indicate that rating scales were used coherently throughout the grid. Thus, the change in predominent ratings did not affect construct variance and therefore probably did not affect the extraction of underlying dimensions.

7.2.3 Response pattern in elements

The statistical tests performed on constructs were also performed on elements. Firstly the pooled rating frequencies of the first nineteen elements on the list were compared with those of the last nineteen elements and the difference was significant ($X^2 = 60.95$, $df = 4$, $p < .001$). A significant difference ($X^2 = 61.77$, $df = 4$, $p < .001$) was also found when the nine elements at each end were compared. The differences between the first nine and the middle twenty elements ($X^2 = 39.87$, df = 4, $p < .001$), and between the middle twenty and the last nine elements ($X^2 = 19,64$, df = 4, $p < .001$) were also significant.

These tests indicated that the first elements tended to receive extreme ratings (1's and 5's); elements in the middle of the list received relatively evenly distributed ratings; elements towards the end of the list elicited more moderate ratings (2's, 3's and 4's). Thus rating scales were not used evenly throughout the elements. It was therefore important to perform comparisons on element variances.

A Mann-Whitney U-test was performed to compare the variances of the first nineteen with those of the last nineteen elements. The variance of the first elements was found to be significantly greater ($U = 86$, $p < .02$). This difference remained when the nine elements at each end were compared ($U = 12$, $p < .02$). However, when the same test was performed to compare each end with the middle twenty elements no significant difference was found.

These results indicate that the shift from extreme ratings to more moderate ones was very gradual. As pointed out earlier, no difference was found in construct variances. Thus the shift in element ratings is likely to have occurred similarly in all constructs. It is expected, therefore, that the change in element ratings should have little effect on the extraction of dimensions since it is based on construct variance rather than on element variance. Care should be taken, however, when interpreting the element contents of emerging dimensions.

7.2.4 Differences between men and women

Of the eleven UG respondents three were women and eight were men. When the ratings of grids from female respondents were pooled and compared to those from male respondents, significant differences were found ($X^2 = 103.51$, df = 4, p <<.001). Women assigned fewer ratings of 2 and 4 and more ratings of 3 and 5. This could indicate a tendency for women to use the rating scale as a three-point scale even if all five points were labelled.

Variances for males and females were compared and no significant difference was found. This indicates that, although men and women used the rating scale differently, this would not be expected to affect the results of principal component analysis or cluster analysis.

7.2.5 Influence of ethnic origin on ratings

It was pointed out earlier that three foreign-born UG respondents had rejected either one or two constructs from their grid on the basis that they did not understand these constructs. The question therefore arose as to whether ethnic origin had any influence on respondents' ratings. Of the UG respondents, five were british-born and six were foreign-born. The pooled rating frequencies of these two groups were compared and a chi-square was calculated. The results and figures are listed in Table 7.9.

These figures indicate very slight differences in ratings of 4. However differences are more noticeable for other ratings. For

TABLE 7.9

A TEST OF THE INFLUENCE OF UG RESPONDENTS' ORIGIN ON RATING FREQUENCIES

		1	2	3	4	5	TOTAL
BRITISH-BORN	OBSERVED	797	1142	859	1264	1068	5130
	EXPECTED	707.3	1313.1	996.6	1232.5	880.5	
FOREIGN-BORN	OBSERVED	738	1708	1304	1411	843	6004
	EXPECTED	827.7	1536.9	1166.4	1442.5	1030.5	
TOTAL		1535	2850	2163	2675	1911	11134

$X^2 = 173.21$

$df = 4$

$p << .001$

british-born respondents the observed frequencies of extreme ratings (1 and 5) were greater than the corresponding expected frequencies. Conversely, the observed frequencies of 2's and 3's were greater than their expected frequencies for foreign-born respondents. It appears that foreign-born respondents tended to be less extreme in their ratings than were british-born respondents. This comparative reluctance to use extreme ratings may have resulted from foreign-born respondents being less sure than british-born respondents of their understanding of constructs.

The total variance of the ratings from foreign-born respondents was found to be somewhat smaller; thus, it is possible that these respondents will account for slightly fewer extracted dimensions. On the other hand, the rating frequencies of both groups were somewhat similar; both distributions can be represented graphically as bi-modal curves whose peaks represent ratings of 2's and 4's. Thus the tendency for foreign-born respondents' grids to yield fewer dimensions may not be noticeable.

7.3 THIRD SAMPLE (SA)

Analysis of response patterns from the third sample proved difficult for three main reasons. Firstly in this sample different respondents used different constructs. Secondly, of the seventeen respondents, five used a five-point rating scale (from 1 to 5), eleven respondents used a six-point rating scale (from 0 to 5), and a seven-point scale (from 1 to 7) was used in the remaining grid.

Thirdly within some grids two or more constructs turned out to be identical, having correlations of 1.0 with each other. As discussed in Chapter 10, only one of such identical constructs was retained for further analyses, the others being discarded.

A summary of rating frequencies on the various rating scales is presented in Table 7.10. Means and standard deviations are also listed for each grid, firstly as a function of the whole rating scale, then as a function of ratings of 1 to 5 only. It must be pointed out that for grid SA14, a seven-point scale is artificially transformed into a five-point scale by equating 2's and 3's to 2's, and 5's and 6's to 4's. Such a procedure, however, may be making undue assumptions. Therefore, in subsequent calculations and comparisons of rating frequencies grid SA14 is left out altogether, although its underlying dimensions are later extracted.

SA respondents were given the opportunity of indicating that a construct was irrelevant to the assessment of an element. They did so by assigning a zero instead of a rating between 1 and 5. Eleven respondents availed themselves of this opportunity, in three cases (SA09, SA10 and SA12) more than 200 times. As the other three samples did not have a similar opportunity, comparing rating frequencies between samples presented a problem.

It can be argued that zero was not a rating as such; thus including in inter-sample comparisons ratings from 1 to 5 from grids in which 0's were used would represent comparing complete grids with parts of grids. Therefore, all subsequent calculations were performed twice: once including and once excluding ratings from grids in which

TABLE 7.10

SUMMARY OF RATING FREQUENCIES, MEAN AND STANDARD DEVIATION FOR INDIVIDUAL SA GRIDS

GRID	0	1	2	3	4	5	6	7	TOTAL	MEAN	S.D.	TOT. 1-5	MEAN	S.D.
SA01	-	27	103	121	84	43	-	-	378	3.03	1.11	378	3.03	1.11
SA02	9	67	62	71	88	60	-	-	357	2.96	1.44	348	3.03	1.37
SA03	-	104	135	123	56	23	-	-	441	2.45	1.13	441	2.45	1.13
SA04	-	153	80	81	58	69	-	-	441	2.57	1.46	441	2.57	1.46
SA05	27	118	91	93	52	81	-	-	462	2.58	1.54	435	2.74	1.45
SA06	-	74	84	91	88	104	-	-	441	3.15	1.41	441	3.15	1.41
SA07	-	78	95	134	127	91	-	-	525	3.11	1.30	525	3.11	1.30
SA08	9	10	49	85	113	49	-	-	315	3.37	1.18	306	3.46	1.04
SA09	304	46	46	2	44	41	-	-	483	1.09	1.71	179	2.93	1.56
SA10	207	41	84	61	71	19	-	-	483	1.60	1.64	276	2.79	1.18
SA11	33	36	107	129	137	41	-	-	483	2.88	1.32	450	3.09	1.10
SA12	268	29	24	51	18	30	-	-	420	1.08	1.64	152	2.97	1.35
SA13	35	92	61	161	52	61	-	-	462	2.62	1.45	427	2.83	1.29
SA14	-	42	68	70	64	45	47	21	357	3.64	1.75	289	3.01	1.29
SA15	73	133	40	71	52	114	-	-	483	2.49	1.83	410	2.94	1.62
SA16	41	110	109	107	95	63	-	-	525	2.56	1.49	484	2.78	1.34
SA17	82	150	107	81	33	9	-	-	462	1.70	1.26	380	2.06	1.08
TOTAL	1088	1310	1345	1532	1232	943	47	21	7518	2.48	1.64	6362	2.87	1.34

o's were found.

As pointed out earlier, only one from each set of perfectly correlated constructs was retained. In all subsequent analyses, the ratings of discarded constructs were removed. The summary of corrected frequencies is listed in Table 7.11. It is on all or parts of the figures of this table that calculations presented in this section were performed.

7.3.1 Uneven ratings

Rating frequencies were compared with an even (or random) distribution in two ways. Firstly, all ratings between 1 and 5 were pooled (except for grid SA14) and these total frequencies were compared to even frequencies. A chi-square of 122.5 was obtained (df = 4, p << .001). Secondly, the ratings from those grids using a five-point scale were pooled. Then, the same test was performed. This time the chi-square was 66.99 (df = 4, p <<.001). Both tests revealed a large number of ratings of 2 and 3 and a small number of 4's and 5's. This suggests that the tendency was present in most grids. However, no explanation can be readily suggested to account for such a phenomenon.

Because different respondents in this sample used different constructs, it is probably invalid to attempt other comparisons by pooling rating frequencies.

7.4 FOURTH SAMPLE (PH)

There were fewer ratings in the whole of the fourth sample than there were in the smallest grid in any other sample. Thus very few comparisons were possible within this sample. The only test which

TABLE 7.11

SUMMARY OF CORRECTED RATING FREQUENCIES, MEAN AND STANDARD DEVIATION FOR INDIVIDUAL SA GRIDS

GRID	1	2	3	4	5	TOTAL	MEAN	S.D.
SA01	27	103	120	70	40	360	2.98	1.11
SA02	67	62	71	88	60	348	3.03	1.37
SA03	104	135	123	56	23	441	2.45	1.13
SA04	153	80	81	58	69	441	2.57	1.46
SA05	118	91	93	52	81	435	2.74	1.45
SA06	74	84	91	88	104	441	3.15	1.41
SA07	78	95	134	127	91	525	3.11	1.30
SA08	10	49	85	113	34	291	3.38	1.01
SA09	46	45	2	44	41	178	2.69	1.59
SA10	37	76	55	61	18	247	2.79	1.18
SA11	23	85	99	108	30	345	3.11	1.08
SA12	29	24	51	18	30	152	2.97	1.35
SA13	88	56	153	49	60	406	2.84	1.30
SA15	118	40	71	52	108	389	2.98	1.60
SA16	110	109	107	95	63	484	2.78	1.34
SA17	78	65	40	16	9	208	2.10	1.12
TOTAL	1160	1199	1376	1095	861	5691	2.88	1.06

could be performed was to check for evenness (or randomness) of rating frequencies. These rating frequencies are shown in Table 7.12. The mean and standard deviation of each individual grid are also listed in the table.

The test was performed on the total frequencies listed at the bottom of Table 7.12. It was found that these were significantly different from a random distribution ($X^2 = 49.6$, df = 4, $p < .001$).

It appears that PH respondents have assigned more ratings of 1 and 2 than ratings of 3, 4 or 5. The phenomenon is somewhat similar to that found in the third sample but involves the other end of the rating scale.

These findings cannot be explained by the grid rating process itself. However, both comparisons between samples and analysis of grid contents did shed some light on these phenomena. These e discussed in the rest of this chapter and in the next three chapters.

7.5 COMPARISONS BETWEEN SAMPLES

It was pointed out in Chapter 5 that the grids used by different samples were different on a number of methodological features. Some of these features are reviewed in this section, and their potential influence on grid rating is assessed.

TABLE 7.12

SUMMARY OF RATING FREQUENCIES, MEAN AND STANDARD DEVIATION

FOR INDIVIDUAL PH GRIDS

GRID	1	2	3	4	5	$\bar{X}$	S.D.
PH01	20	21	3	7	3	2.11	1.20
PH02	23	11	3	13	4	2.33	1.41
PH03	18	13	8	5	10	2.56	1.49
TOTAL	61	45	14	25	17	2.33	1.38

7.5.1 Effects of labelling

One of the features on which grids were different was the number of points of the rating scale which were labelled. For instance, GR and PH respondents used rating scales labelled only at both ends. The rating scale used by SA respondents was labelled at mid-point as well as at both ends. Finally, all points of the UG scales were labelled.

It was argued earlier that there could be a tendency for respondents to use mostly labelled ratings. It was also argued that labelling all points of a scale might facilitate the use of intermediate ratings (2's and 4's). This was tested by comparing samples as a function of their number of labelled scale points.

For the purpose of this calculation GR and PH frequencies were pooled. The calculation was performed twice. For both tests, ratings of rejected constructs were excluded from GR and SA frequencies. The first calculation included all corrected SA ratings between 1 and 5 except those from grid SA14; the second calculation included corrected ratings only from those SA grids with a five-point scale.

The figures used in the first calculation, along with its results, are shown in Table 7.13. The table shows that a chi-square of 1472.7 was calculated (df = 8, $p << .001$).

Although the figures from the second calculation are not tabulated in this section, the results were nearly identical ($X^2 = 1438.2$, df = 8, $p << .001$). Furthermore, the same pattern of differences was observed

TABLE 7.13

A CHI-SQUARE TEST OF THE EFFECTS OF LABELLING ON RATING FREQUENCIES

		1	2	3	4	5	TOTAL
*GR & PH	O	1420	463	720	369	952	3924
	E	778.2	853.3	805.5	782.8	704.3	
**SA	O	1160	1199	1376	1095	861	5691
	E	1128.7	1237.5	1168.2	1135.2	1021.4	
UG	O	1535	2850	2163	2675	1911	11134
	E	2208.1	2421.2	2285.4	2221	1998.3	
TOTAL		4115	4512	4259	4139	3724	20749

$X^2 = 1472.7$

$df = 8$

$p \ll .001$

* Excluding ratings on rejected constructs.

**Excluding O's and ratings from grid SA14.

in the second calculation as that which can be seen from the figures of the first calculation.

Table 7.13 shows that, except for ratings of 5 in SA grids, all labelled points in GR, PH and SA grids show observed frequencies greater than expected frequencies. Conversely, all unlabelled points show observed frequencies smaller than expected frequencies. In UG grids, where all points were labelled, intermediate ratings (2's and 4's) show observed frequencies greater than expected frequencies.

These findings suggest that labelling the points of a rating scale facilitates the use of the labelled points. These indications were further checked by comparing individual grid variances between groups. It was expected that the more points that were labelled on a rating scale the smaller the variance of the grids with such scales would be.

It was found that the variance of GR grids was significantly greater than that of grids from any other sample, with probabilities at least smaller than .025. Other samples did not show significant differences of variance. Thus, although there were indications that labelling influenced the way in which rating scales were used, another factor, discussed hereafter, also had an influence upon the rating process.

7.5.2 Cognitive complexity

Bieri (1955) argues that the complexity of a person's thought processes plays a major role in his construing of reality. It is possible that higher education can increase this complexity by training a person to discriminate better between alternatives. Greater complexity may have led to a more even use of rating scales by respondents with university education.

Respondents' imputed cognitive complexity therefore might be expected to account, at least in part, for the results mentioned above. Only one GR respondent had received university training. In contrast, all respondents in the other samples were receiving or had received university education.

An indication that cognitive complexity may have influenced ratings was found earlier in the comparisons of grid variances be en samples. Although both GR and PH samples used rating scales labelled only at both ends, GR variances were significantly ($p < .025$) greater than PH variances.

The rating frequencies of these two samples were compared. A chi-square was calculated to test for differences between the two sets of figures. These figures and the results of the test are shown in Table 7.14. The table shows that PH respondents assigned more ratings of 2 and 4 than expected. Conversely, there were more ratings of 3 and 5 than expected in GR grids. It thus appears that imputed cognitive complexity was associated with a more discriminating use of rating scales.

TABLE 7.14

A TEST OF THE INFLUENCE OF COGNITIVE COMPLEXITY ON THE USE OF RATING SCALES

Rating	PH Observed	PH Expected	GR* Observed	GR* Expected	TOTAL
1	61	58.6	1359	1361.4	1420
2	45	19.1	418	443.9	463
3	14	29.7	706	690.3	720
4	25	15.2	344	353.8	369
5	17	39.3	935	912.7	952
TOTAL	162		3762		3924

$X^2 = 65.18$

$df = 4$

$p < .001$

*Corrected frequencies, excluding frequencies from rejected constructs.

7.6 SUMMARY

In this chapter, a number of variables which influenced grid rating were identified. For instance, it was found that a respondent's job, sex and ethnic origin appear to affect the way in which respondents use rating scales.

Differences were also found between respondents with and without university training. Further differences appear to have been a function of the number of rating scale points which were labelled. Finally in very large grids used by UG respondents, rating patterns appeared in the latter parts of constructs and towards the end of the grid.

Some possible influences of these rating differences on the emergence of underlying dimensions were postulated. It is possible that some of these influences are mutually enhancing. It is also possible that certain influences counteract others. Thus, it is difficult to predict accurately what their resulting effect is. However, it is important to bear these rating differences in mind when analysing hazard assessment patterns.

CHAPTER 8 : RESULTS FROM THE FIRST SAMPLE

The first step in the identification of hazard assessment dimensions - described in Chapter 5 - was an analysis of individual grids. A comparison of results from individual grids within samples was the second step, and the final step was to compare consensus dimensions across samples.

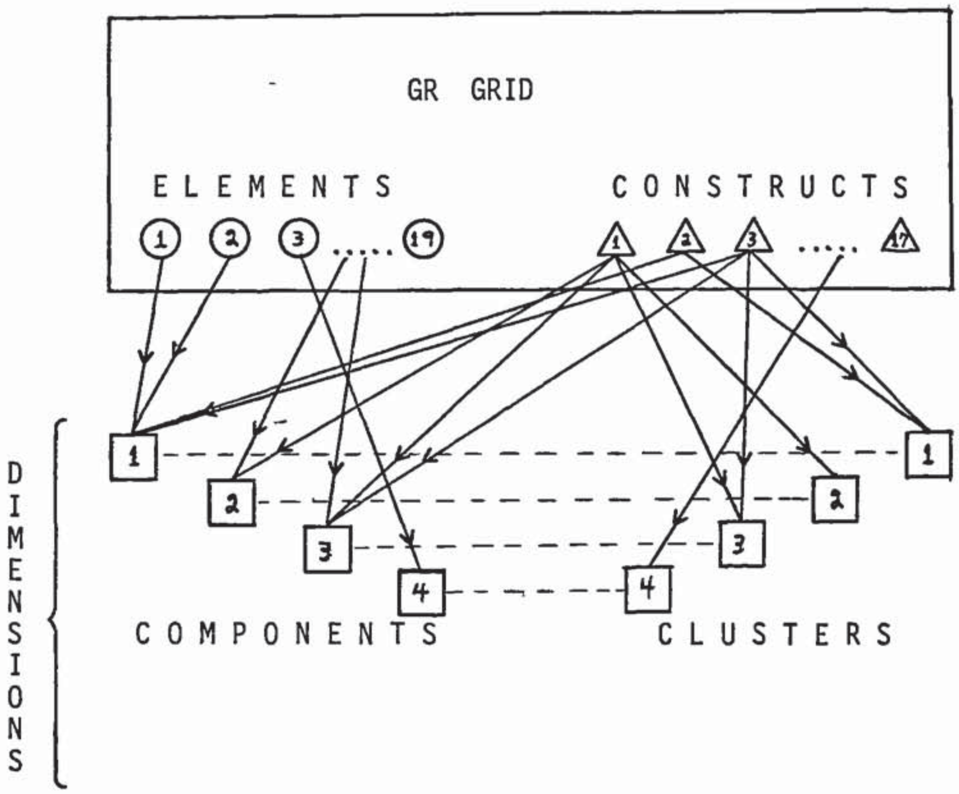
This chapter is a discussion of the first two types of analysis on the data provided by GR respondents. Firstly, the results of the analysis of individual grids are presented; these results also shed some light on certain methodological flaws in the analyses which need to be kept in mind in subsequent chapters. Secondly, the consensus dimensions are similarly discussed. Then, the labels affixed to the various dimensions identified in this sample are compared to the predicted labels. A further discussion of these results can be found in Chapter 11, where comparisons are made between samples.

8.1 INDIVIDUAL GRIDS

In the analysis of individual grids, principal component analysis and Auclair's cluster analysis were performed on constructs. Elements were subjected only to principal component analysis. Figure 8.1 is a diagram which may help to visualize how dimensions are extracted, i.e. components from elements and constructs, and clusters from constructs.

FIGURE 8.1

DIAGRAM OF THE RELATIONSHIPS BETWEEN THE CATEGORIES OF ITEMS



The results of these operations on constructs are summarized in Table 8.1. The outcome of each grid can be read vertically, and what happened to each construct can be read horizontally. The abbreviations Q and F refer to cluster and to component respectively; the numbers in the columns are the rank (first, second, etc.) of the extracted dimension (cluster or component) to which the construct belonged. This convention will be used throughout the analyses. As for construct and element numbers, these refer to numbers used in Table 4.6 and 4.5 respectively. For reading convenience these tables can be found in fold out form in Appendices 1 and 2.

8.1.1 Number of components

In Table 8.1, NQ is the number of clusters and NF is the number of components in each grid. Whereas determining the number of significant clusters was relatively straightforward, a problem arose over the number of components extracted from each grid. The Bartlett test gave negative results for four grids (GR05, GR10, GR11 and GR12); in other words no component was found significant in these grids. Only two components were found significant in grid GR07.

Unless instructed otherwise Ingrid 72 lists the details of only those components which are found significant. However, because the main hypotheses of this research postulated the existence of three important dimensions of hazard assessment, it would be difficult to fully discuss three dimensions if details of one or more of them were missing from individual grids. More comprehensive conclusions could be expected to be reached about the postulated dimensions if full details of these could be studied in all individual grids. Ingrid 72

TABLE 8.1

SUMMARY OF THE RESULTS OF CLUSTER ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS ON CONSTRUCTS IN GR GRIDS

CONSTRUCT	GR01		GR02		GR03		GR04		GR05		GR06		GR07		GR08		GR09		GR10		GR11		GR12	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F
1	5	3-4	1	1	1	1	3	1	3	3	1	1	1	1			1	1-2	1	1				
2	1	1	5	1	3		1	1	2	1	4	1-3				1		3	1	1	1	1		2
3	1	1	2	4	3	3-4	3	1-2		1-2	1	1	3	1-2	3	2		3	1	1		2		1-2
4	5	2-4		3		1		3			2	2		2	1	1	2	2		3	X	X		
5		6	4	2		3	1	1		3		4	4		1	1-2	2	1				1		
6	1	1	1	1-3	1	1	1	1		1	3	3		3	2	2	1	1	3	2			3	1-3
7	1	1		3	1	1	1	1	2	1	3	2-3	2	3	1	1				2	X	X	4	1-2
8	3	1-4	1	1	4	3	3	2-3	3	1	2	2	4	1-3	3	2	3		1	1		3	1	1
9	1	1	4	2	1	1	4	4			3	3	2	1-3	2	2	2	1-2	X	X	1	2	4	1
10		1	3	.	1	1	4	3-4		1	2	2	2	3	2	2	2	2		1	2	1-2	2	2
11	4	5	5	1	1	1	3	3-4	3	3	2	2	5		2	1-2	3	3					3	3
12	X	X	4	2	2	1-2	1	1	X	X	1	1	4	2		3	1	1	1	1	X	X	2	2
13	2	2	5	1	1	1	2	2-3	1	2	1	2	3	1-2	1	1	1	1-2	2	2	1	2	3	1
14	3	1-2	2	1-4	2	2	2	2	1	2	1	1	5	2		1	1	1	1	1		3	1	1
15	2	1-2	2	1	2	2-3	2	1-2	1	2-3	4	1	1	1-2	1	1			2			3	4	1
16	2	2	3	1	2	2	2	2	1	2	4	1	1	1	1	1	1	1	1	1-3	2	3	1	1
17	4	3	3	1-3	4	2		6	2	1-2		4	1	1	1	1	1	1	3	2-3	2	1	1	2
NQ	5		5		4		4		3		4		5		3		3		3		2		4	
NF		6		4		4		6		3		5		3		3		3		3		3		3

was therefore instructed to provide full details of at least three components. Following this decision, in interpreting results allowance will have to be made for the fact that some dimensions in certain grids were found to be not significant.

In Table 8.1, the NF figures refer to a minimum of three components for five of the twelve grids. For these five grids, the details of constructs assigned to components were tabulated as if three components had been found to be significant. For the other seven grids, all those components found to be significant are listed.

The fact that not one component was found to be significant in four of the twelve grids could be attributable to one of two reasons. The first is mentioned by Slater (1972) in a discussion of possible inadequacies of the Bartlett test. The second reason may be that grids which yielded no significant component were scored by respondents whose patterns of hazard assessment were loosely structured; such patterns would not be expected to show an underlying trend. It was decided to investigate both possibilities.

8.1.2 Inadequacies of the Bartlett test

Slater (1972) writes that in some circumstances the results of the Bartlett test can be misleading. According to Slater (1972) in some grids variance on some constructs is unduly restricted, which in turn yields low latent roots. In contrast, low variances can sometimes be associated with unusually high variances within the same grids. Very high and very low variances in turn influence the significance of components.

Low variance leads to low latent roots; but roots and correlations being interdependent low variance usually accounts also for low correlations. Mann-Whitney U tests were therefore calculated to compare highest correlations between grids which yielded less than three significant components and grids which yielded three or more significant components. The same operation was performed by using latent roots of first components instead of highest correlations.

The results of the comparison of highest correlations are shown in Table 8.2. It can be seen from the Table that the highest correlation for GR10 was 1.000. The correlation was negative and was found between constructs 2 and 12; this means that respondent GR10 considered that the more a hazard was a necessary result of a process the less preventable it was. For statistical purposes, the two constructs are taken to measure the same thing, albeit in reverse. The calculations were therefore performed twice more: one using the second highest correlation for GR10, and once excluding GR10 altogether. Whereas the original calculation was found to be barely significant ($p = .074$), the second and third were found to be significant ($p = .024$ and $p = .006$ respectively). Highest correlations therefore seem to have been significantly lower for grids in which less than three components were found significant. Low correlations therefore appear to have affected the outcome of the Bartlett test.

It was also postulated that low latent roots might explain why the Bartlett test found few if any significant components in some grids. If the hypothesis was right, the highest root (that of the first component) would tend to be lower in those grids which yielded few significant components. This was tested, and the results are shown in Table 8.3.

TABLE 8.2

TEST OF THE INFLUENCE OF LOW CORRELATIONS ON THE OUTCOME OF THE BARTLETT TEST

Grid	Highest correlation in absolute value	Rank: Less than 3 significant components		Rank: 3 or more significant components
GR01	.950			11
**				(12)
GR02	.729			4
GR03	.831			8
GR04	.942			10
**				(11)
GR05	.732	5		
GR06	.797			6
GR07	.718	3		
GR08	.815			7
GR09	.856			9
**				(10)
GR10*	1.000	12		
**	(.855)	(9)		
GR11	.714	2		
GR12	.667	1		
	N =	5		7
	Sum of ranks =	23		55
	U Statistic =	27	p = .074	8
**	Sum of ranks =	20		58
	U Statistic =	30	p = .024	5
*	N =	4		7
	Sum of ranks =	11		55
	U Statistic =	27	p = .006	1

** : A second calculation was performed using the second highest correlation in GR10 and adjusting ranks accordingly.

* : A third calculation was performed excluding GR10.

TABLE 8.3

TEST OF THE INFLUENCE OF LATENT ROOTS ON THE OUTCOME OF THE BARTLETT TEST

Grid	Root of first component	Rank: Less than 3 significant components	Rank: 3 or more significant components
GR01	5.437		11
*			(10)
GR02	4.085		4
GR03	4.994		8
GR04	4.710		7
GR05	3.565	2	
GR06	4.523		6
GR07	4.086	5	
GR08	5.383		10
*			(9)
GR09	5.462		12
*			(11)
GR10*	5.299	9	
GR11	2.781	1	
GR12	3.974	3	
	N =	5	7
	Sum of ranks =	20	58
	U Statistic =	30 p = .024	5
*	N =	4	7
	Sum of ranks =	11	55
	U Statistic =	27 p = .006	1

* Calculation excluding grid GR10

There proved to be a significant ($p = .024$) difference between latent roots; latent roots of first components tended to be lower in those grids where fewer or no significant components were extracted. For grid GR10, not only was there no significant component, but the Bartlett test was not even applied. This indicated that Ingrid 72 had detected something wrong with variances in this grid. It is thought that the correlation of 1.00 may have had something to do with the test not being applied. The difference in latent roots was therefore calculated a second time, excluding GR10 from the calculation. The difference proved to be even more significant ($p = .006$).

8.1.3 Cognitive structure

On the basis of the discussion above the results of the Bartlett test cannot be accepted unquestioningly for the five grids from which less than three significant components were extracted. There are further indications which cast some doubt about the results of the Bartlett test on these five grids.

For instance, Slater (1972) points out that low variance on some constructs can sometimes be accompanied by very high variance on other constructs within the same grid. This seems to have happened in grids GR11 and GR12; whereas highest correlations and latent roots for those two grids were among the lowest of all individual grids their mean construct variation were the two highest. Finally, Slater (1972) points out that the Bartlett test can be unreliable when a respondent has failed to discriminate between two constructs; such a failure has occurred in grid GR10, when constructs 2 and 12 correlated at -1.00.

Indications were needed, however, as to whether analysis of at least three dimensions was justified for the five grids for which the Bartlett test seems to have been inadequate, or whether the five grids discussed so far were scored by respondents whose hazard assessment patterns were loosely structured. In pursuing this point, Cattell's Scree test and Kaiser's criterion of a latent root of 1 - described in Chapter 5 - were applied to the five grids. In addition, the numbers of significant clusters were examined. The three sets of figures are listed in Table 8.4.

Because of limitations in all three methods described above, none of them can be assumed to be more reliable than the others. For instance, none of the Scree test graphs showed a neat two-segment curve as described by Cattell (1953) and by Child (1976); all graphs showed at least three or four noticeable changes in slope. So the results indicated in Table 8.4 refer to the component at which the first (and usually the most marked) slope change occured in each graph. As for Kaiser's criterion, it was pointed out earlier that a latent root of 1 can often be too lax. Finally, because it used each test item only once, cluster analysis often yields fewer dimensions than does factor analysis.

These arguments are illustrated in Table 8.4. Thus the largest number of dimensions is always obtained by the latent root of 1. Conversely, the smallest number of dimensions is always, except for GR07, the number of clusters. Only two clusters were extracted from GR11; but both indicators of the number of components in the same grid agree on more than three dimensions. In the other four grids, even cluster analysis reveals at least three dimensions.

TABLE 8.4

INDICATIONS OF THE NUMBER OF SIGNIFICANT DIMENSIONS

IN GRIDS WHICH YIELDED LESS THAN THREE COMPONENTS

Grid	Scree Test	Latent Root of 1	Number of Clusters
GR05	5	6	3
GR07	4	7	5
GR10	5	6	3
GR11	5	6	2
GR12	-5	6	4

On the basis of these indications it would appear that failure by the Bartlett test to retain three significant components was due to inadequacies in the test rather than to respondents' loose cognitive structure. Thus, it was decided, for further analyses, to include the particulars of three components for those grids in which there are many indications that the Bartlett test was unreliable.

The Bartlett test did prove to be somewhat inadequate for five of the twelve grids, because of such inherent characteristics of the grids as: abnormal variance, low correlations and latent roots, and failures to discriminate between constructs. There is an additional indication that probelms in extracting dimensions seem to stem from the respondents themselves; of the five problematic grids, three (GR05, GR10 and GR11) contained at least one rejected construct. Allowances will have to be made for peculiarities in those grids when further analyses are discussed.

8.1.4 Constructs and elements

As mentioned in Chapter 7, some constructs were rejected because respondents rated all elements identically on them. These constructs are indicated by X's in Table 8.1.

Another noticeable feature of Table 8.1 is that in some grids, a number of constructs were either not assigned to any cluster or to any component; there are also instances where a construct was assigned neither to a cluster nor to a component. In all, 204 constructs could be assigned (17 constructs from 12 grids). 9.8% of those were not assigned to any component, and 22.1% were not assigned to any cluster. In other words, 15.9% of the cells in Table

8.1 are empty.

It is not surprising that there were fewer assignments to clusters than to components. As pointed out earlier, cluster analysis is more stringent a technique than principal component analyses. What is somewhat surprising is that the results of both analysis are less in agreement than expected.

As pointed out in Chapter 5 and at the beginning of this chapter, principal component analysis was also performed on elements. A summary of the results of this analysis appears in Table 8.5. This table shows that, out of a possible 228 elements (19 elements from 12 grids) 41.7% were not assigned to any component. This figure is even greater than the percentages of unassigned constructs. Furthermore, a first glance at the assigned elements seems to reveal less unanimity than was found in constructs.

This apparent lack of unanimity in assigned constructs and elements can be the symptom of two types of problems. Firstly, it is possible that lack of unanimity may reveal a lack of robustness in extracted dimensions. It could also be that lack of unanimity stems from non-comparable dimensions between grids. Both possibilities were looked into.

8.1.5 Robustness of dimensions

A dimension is robust if its identification or extraction can be repeated. Two methods of extraction were retained for this study. Thus dimensions can be said to be robust if they are identified by both methods.

TABLE 8.5

SUMMARY OF OUTCOME OF PRINCIPAL COMPONENT ANALYSIS ON ELEMENTS

ELEMENT	GRID											
	GR01	GR02	GR03	GR04	GR05	GR06	GR07	GR08	GR09	GR10	GR11	GR12
1	2	1		1-2	1	1		2	1	1		
2	1		1-2	2	5		1	1	1-2			1
3	1	2	1	1	3	3	3	2	3			3
4	1	1-4	2-4	1	2	1			2		2	1-2
5	2	2-4	3	1					2			1
6	1	1			1-2			3		1		
7	1			2		2	2	1	2	2	2	1
8	1	2		2		2	1	1	1	1		
9				6			1	1				
10	1		3	4			1					
11	1	3	1			2-3-4		3	2		2	
12				1-3		1	2	2		1	3	1-2
13		3		3	2	2-3	2-3	3	2	1		
14		3	2	2-3-5	2	1		1-2		1		
15		2	1-4	3	1-3	1	1-2-3	1-2	1			1
16	3	1-2				2						
17	2	1	1		1	4		1	1-2		1	
18	2		1	3	2		1	1	1	2	1	
19	2		1		3				2	2		

Some of the discrepancies found in Table 8.1 are differences in dimensions to which a construct was assigned within an individual grid. For instance, construct 1 in grid GR01 was assigned to cluster number 5 and to components number 3 and 4.

The findings in Table 8.1 were scrutinized thoroughly in order to check whether within a grid the same dimensions were extracted by both analyses. A number of points came to light following this scrutiny.

Firstly, it was found that more than half the components could be identified to over three quarters of the clusters. For instance, throughout the twelve grids, fifteen components (out of forty-six) could be said to encompass approximately the same constructs as the corresponding cluster. Four components were found to cover the corresponding cluster along with one subsequent cluster. Two more components were found to encompass the corresponding cluster along with two other clusters. Three more components were found to be similar to a cluster which did not have the same rank. Finally, two components were found to have the same core as the corresponding cluster.

This meant that twenty-six components were matched, possibly with slight differences, to thirty-three clusters. The slight differences found were, in the vast majority of cases, a greater number of constructs in components than in clusters. The additional constructs were for the most part the last one retained in the components.

These differences, along with the breakdown of some components into two or more clusters, might cast some doubt about the robustness of dimensions. However one must remember that cut-off points tend to be higher in cluster analysis than in component analysis, thus making cluster analysis a more stringent method. Higher cut-off points can easily explain why clusters included fewer constructs and why groups of clusters were pooled into single components.

There remained twenty components (out of the forty-six) and twelve clusters (out of forty-five) which could not be matched. It must be pointed out that, of these unmatched dimensions, only seven were prominent (first or second) clusters or dimensions; and six of the seven were found in the highly problematic grids GR10 and GR11. Three more unmatched dimensions are also found in those two grids.

Therefore, the prominent (first and second) dimensions tended to be robust. Problems occurred mostly with third and subsequent dimensions.

Of the unmatched components six encompassed only one construct and very few (if any) elements. Such components would probably be impossible to label.

Of the unmatched clusters, eight included only two constructs. It is very likely that those eight clusters were by-products of the method chosen. For instance, when most items in a matrix have been assigned to a cluster, two of the remaining items may form a cluster if the correlation between the two is significant. And yet, this correlation may be due for instance to a cause-and-effect relationship between the items rather than to a common trait in both items.

One final point deserves a comment about the results in Table 8.1. There were twelve grids in which a construct could be assigned either to a cluster, to a component, or to both; this represented a maximum of twenty-four possibilities of assignment. Constructs 4 and 5 were rejected or failed to be assigned for more than a third (eight) of those possibilities. Four more constructs (number 1, 2, 7 and 12) were rejected or failed to be assigned for a quarter (six) or more of those possibilities. This may mean that those particular constructs need to be treated carefully. They are discussed at greater length later.

To summarize, a greater proportion of clusters than of components tended to be robust. But in general, the prominent dimensions were replicated by both methods of analysis, even if it meant that two or more clusters were covered by one component. Quite a few of the differences found in matched dimensions are most probably attributable to cluster analysis being a more stringent method.

8.1.6 Comparability of dimensions

Two methods were used in an attempt to assess the comparability of dimensions between grids. The first method was to affix labels to all dimensions in all grids in order to check whether similar labels could be found across the grids.

Some problems arose when this method was tried. It was found that quite a few dimensions were very difficult to label. This difficulty tended to be greater for clusters than for components; components generally had more constructs and also included elements. But even components sometimes presented labelling problems, as a large number of them were still relatively small. Twenty-seven of the forty-six components encompassed less than a quarter (nine) of the thirty-six items (seventeen constructs and nineteen elements).

Thus, it was decided not to pursue the labelling exercise further on this sample. The second method of comparison was carried out instead. This consisted of a search for cores of constructs which might be present in one of the dimensions in most grids. But the results of this search were hardly more satisfactory than those of the first method.

For instance, no core of three constructs could be foun to appear in at least one dimension in all grids. Only constructs 14 and 16 were associated within a dimension in eleven grids. Constructs 14 and 15 were paired in ten grids, and so were constructs 15 and 16. However, the triad 14-15-16 (the most frequent one) was found in only six grids. Constructs 13 and 17 were often associated individually to each construct of the triad; but when they were added to the triad, the five constructs were associated in only one grid.

What this seems to indicate is that cut-off points in both cluster and component analysis may have been somewhat high. For instance, the lowest cut-off points in principal component analysis is 0.372. In comparison for UG grids, the corresponding figures were

0.325 and 0.265 respectively.

High cut-off points would explain why constructs paired in a majority of grids seldom formed a solid core. They would also explain why GR dimensions generally encompassed a smaller proportion of the total number of constructs than did UG dimensions. Finally, high cut-off points could also explain why some components were fragmented into two or more clusters.

As they stand, the results tend to invalidate the hypothesis that hazard assessment patterns exist which are used by most people. However, in view of the problems probably caused by high cut-off points, it was decided to proceed with the analysis of consensus dimensions. It was hoped that this analysis might provide indications about the answer to the problems outlined above.

8.2. CONSENSUS DIMENSIONS

As described in Chapter 5, three techniques were used to extract underlying dimensions: an analysis of frequencies based on Table 8.1, a grid of mean cell scores ($\overline{X}$), and a grid of mean Z-scores ($\overline{Z}$). The outcome of those three analyses is summarised in Table 8.6.

For each construct Table 8.6 lists the number of individual grids in which a construct was found to have been assigned to a dimension by principal component analysis and by cluster analysis. The highest frequencies are underlined for each construct and for each method. Table 8.6 also describes the dimensions to which constructs were assigned by the two statistical analyses on both the $\overline{X}$ grid and the $\overline{Z}$ grid.

TABLE 8.6

SUMMARY OF RESULTS FOR THE THREE METHODS OF EXTRACTING CONSENSUS DIMENSIONS

Construct	DIMENSIONS 1		2		3		4		5		6		None		Rejected	$\overline{X}$		$\overline{Z}$	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F		Q	F	Q	F
1	6	7		1	2	2		1	1				3	3		3	2-3	3	2-3
2	4	8	1	1	1	2	1		1				4	2		1	1	1	1
3	3	7	1	6	4	2		2					4			3	1-3	3	1-3
4	1	2	2	4		3		1	1				7	2	1	2	2	2	2
5	2	4	1	2		2	2	1				1	7	3		3	3	3	3
6	5	7	1	2	3	4							3	1		4	4	4	3-4
7	4	6	2	3	1	3	1						3	1	1	4	4	4	4
8	3	6	1	3	5	4	2	1					1	1		1	1	1	1
9	3	5	3	4	1	2	3	1					1	1	1	2	2-3	2	2-3
10	1	5	6	5	1	2	1	1					3	1		2	2	2	2
11	1	3	2	2	4	4	1	1	2	1			2	3		4	2-4	4	2-4
12	4	5	2	4		1	2						1		3	1	1	1	1
13	6	6	3	8	2	1			1							1	1-2	1	1-2
14	5	7	3	5	1	1		1	1				2			1	1	1	1
15	3	7	5	5		3	2						2	2		1	1-4	1	1-4
16	6	7	4	4	1	2	1									1	1	1	1
17	4	6	2	4	2	3	2	1		1			2			1	1	1	1

Probably the most noticeable feature of Table 8.6 is the very great similarity between the results of the $\overline{X}$ grid and those of the $\overline{Z}$ grid. But also remarkable is the similarity between the outcome of cluster analysis and that of principal component analysis in both grids. The latter similarity may not be as obvious as the former, because principal component analysis can assign a construct to more than one component. But, when a construct does appear in more than one component, there is always one of the component numbers which is identical to the cluster number for the same construct.

There is also a fair amount of consistency between principal component analysis and cluster analysis in the summary of frequencies. For example, there are twelve constructs out of seventeen for which highest frequencies occur in the same dimension. Furthermore, twenty-six of the thirty-four frequencies corresponding to the dimensions retained in the $\overline{X}$ and $\overline{Z}$ grids were either the highest or the second highest for the construct. There are, however, constructs which show discrepancies both between frequencies and mean grids, and within frequencies.

Two main explanations can be suggested as to why those discrepancies exist. Firstly, it may be that, as both $\overline{X}$ and $\overline{Z}$ grids are derived from mean scores, they gloss over individual differences and make consensus dimensions look more robust than they are. Alternatively, it may be that the analysis of frequencies is a less powerful technique than the other two. Both possible explanations deserve further consideration.

8.2.1 Further analysis of frequencies

It was clear from Table 8.1 that extracted dimensions varied from one grid to another. Subsequent analyses revealed peculiarities (mostly related to rating frequencies and variance) in five of the twelve grids. These analyses indicate that grids of means may be at least partially unreliable because of important oversights. A further analysis of Table 8.1 was therefore undertaken with reference to the figures in Table 8.6.

The first stage of this re-analysis investigated those constructs for which there was disagreement between methods. They were numbers 1, 5, 6, 7, 8, 9, 11 and 15. It can be postulated that in general, those constructs were subject to less variation; smaller variation usually makes for lower correlations and thus for reduced likelihood of a construct being assigned to a dimension.

The hypothesis that ambiguities in constructs stemmed from low variance was tested by a Mann-Whitney U-test on the standard deviations of constructs. The results of the test are shown in Table 8.7. The standard deviations which led to these results were calculated by excluding the scores of rejected constructs (cf. Table 7.1); but the results of the test were the same when uncorrected standard deviations were used. The variance of those constructs which led to ambiguities was found to be significantly ($p < .05$) lower. This may prove to be important in the interpretation of consensus dimensions; and yet, the grids of means yielded no clue about the problems caused by low variance.

TABLE 8.7

TEST OF THE INFLUENCE OF CONSTRUCT VARIANCE ON DIMENSION AMBIGUITIES

Construct	Standard deviation	Rank: Differences between methods	Rank: No difference
1	1.58	11	
2	1.56		10
3	1.65		14.5
4	1.81		17
5	1.34	6	
6	1.49	8	
7	1.32	5	
8	1.28	4	
9	1.22	1.5	
10	1.27		3
11	1.40	7	
12	1.22		1.5
13	1.69		16
14	1.59		12
15	1.51	9	
16	1.65		14.5
17	1.63		13
	N =	8	9
	Sum of ranks =	51.5	101.5
	U Statistic =	56.5	15.5

$p < .05$

On the other hand, there are also indications that the method of frequencies may have been less robust than the other two methods. A re-analysis of Table 8.1 illustrates this point. For instance, in earlier discussions, it was argued that the notion of preventability was strongly associated with the "controllability" dimension. It was hypothesized that "controllability" would emerge as the most important dimension of hazard assessment.

Construct 12, which measures perceived preventability, was rejected from three grids. It is interesting to note that in those three grids, the "controllability" dimension (component 1) seems less coherent than in other grids such as GR08. In grid GR01, component number 1 is broken into clusters number 1 and 3. In grids GR05 and GR11 component number 1 becomes cluster number 2 and component number 2, cluster number 1. It appears that the absence of a salient concept in three grids may have created confusion in the dimensions which were extracted from those grids. However, confusion also existed in some of the other nine grids. Therefore, the absence of construct 12 is not the only explanation.

Grid variance also seems to have been responsible for some confusion in extracted dimensions. Discrepancies seem to exist in those grids which account for the highest or the lowest grid variance. The high grid variances, it was pointed out earlier, were artificially inflated because of respondents using mostly extreme ratings; this tends to distort the importance of some correlations. By contrast, low variance is difficult to apportion; this may prevent some dimensions, which under different circumstances would be found significant, from being extracted.

Another indication that variance did cause ambiguities emerges when one looks into the dimensions which account for the greatest number of discrepancies. In general, cluster analysis and principal component analysis seem in close agreement over dimension 1. Dimension 2 shows more discrepancies between the two analyses but it is still recognizable in a majority of grids. Discrepancies are more important in dimension 3; there are many constructs which are assigned to it by only one of the two analyses. Finally, only in grids GR01 and GR04 is there an indication of a relatively robust fourth dimension; but in grid GR01, the fourth component corresponds to the fifth cluster. Therefore, it appears that the greater the variance that is explained by a dimension, the more comparable it is between individual grids.

Low variance constructs led to confusing results in the analysis of frequencies. Of the three extraction methods, only the analysis of frequencies could identify the problem of low variance. The frequencies method could not take variance problems into account in identifying underlying dimensions accurately. Pooling the data, however, made it possible to define precisely those dimensions which tended to be present in most grids but which were not quite significant. In that respect, the two grids of means were more robust. But before interpreting the consensus dimensions extracted by the three methods, and in order to facilitate interpretation, the elements were analysed.

8.2.2 Analysis of the elements

When Ingrid 72 was described, it was pointed out that principal component analysis was performed on elements as well as on constructs. The outcome of this analysis on individual grid is summarized in Table 8.5. As Ingrid 72 does not provide the matrix of correlations between elements, cluster analysis was not performed.

The main purpose of the analysis of elements was to enhance or to clarify the meaning of dimensions obtained from constructs. As such, the analysis of elements did not need to be as thorough as the analysis of constructs. Hence, no action was taken to perform cluster analysis. But the three methods of extracting consensus dimensions were applied to elements. These results are shown in Table 8.8.

As in the analysis of constructs, Table 8.8 shows discrepancies between the frequencies and the two grids of means. There is, however, great consistency between the $\overline{X}$ and the $\overline{Z}$ grids. In the light of earlier discussion, it is likely that phenomena related to grid rating and variance influenced the outcome of principal component analysis on elements just as they influenced the analysis of constructs. However, the information provided by Ingrid 72 does not lend itself to further analyses along those lines.

8.3 INTERPRETATION OF CONSENSUS DIMENSIONS

Having extracted dimensions which represented a certain consensus between the individual grids, the next stage was to scrutinize the constructs and elements contained in those dimensions. But in order to do so, it was necessary to assess the relative

TABLE 8.8

SUMMARY OF RESULTS FOR THE THREE METHODS OF EXTRACTING CONSENSUS DIMENSIONS FROM ELEMENTS

ELEMENT	DIMENSIONS 1	2	3	4	5	6	NONE	$\bar{X}$	$\bar{Z}$
1	6	3					4	1	1
2	6	4					4	1	1
3	3	2	5				2	4	4
4	5	5		2			3	1-2	2
5	2	3	1	1			6	1	1
6	4	1	1				7		
7	3	6					3	1-4	1-4
8	5	3					4	1	1
9	2					1	9		
10	2		1	1			8	2-4	2-4
11	2	3	3	1			5	2-3	2-3
12	4	3	2				5	1-2	1-2
13	1	4	5				4	3	3
14	3	4	2		1		5	1	1
15	7	3	3	1			3	2	2
16	1	2	1				9		
17	6	2		1			4	1	1
18	5	3	1				3	1-3	1-3
19	1	3	1				7		

importance of the constructs and elements within the dimensions and to know the sign of the correlations between them.

Indications of the polarity (sign of correlations) could be obtained from the correlations between constructs in cluster analysis and from the sign of loadings both for elements and for constructs in principal component analysis. There were many sources of indications about the relative importance of items within dimensions. For instance, the importance of an element could be judged by its loading in the $\overline{X}$ and $\overline{Z}$ grids and by the number of times it was assigned to a component in individual grids. In addition to these parameters, constructs could also be assessed by the order in which they were included in clusters in the $\overline{X}$ and $\overline{Z}$ grids and by the number of times they were members of certain clusters in individual grids.

It was decided to take as much of this information as possible into account in interpreting the consensus dimensions. Therefore, each type of information was converted into ranks, and items were arranged on the basis of their mean rank. As a result of this task, three minor points deserve attention. Firstly, a cluster is always opened by the highest correlation between two constructs; so the first two constructs were both assigned a rank of 1. Secondly, when a construct was assigned to an important cluster, it was not re-used for less important dimensions; because of this limitation of the technique, the mean rank was calculated on fewer indices when a construct was assigned to a larger cluster. Finally, the results of cluster analysis proved to be identical in both the $\overline{X}$ and the $\overline{Z}$ grids; those results were then considered as one indicator instead of two.

8.3.1 Controllability

Champion (1977) used the label "scope for personal control" to describe the first consensus dimension extracted from the GR grids. But because only the grid of mean cell scores was used, some doubt remained about the accuracy of the label. The same grids have been analyzed in more detail. It is therefore possible to discuss the same dimension more fully.

The details of the first dimension which was extracted from GR grids are listed in Table 8.9. Constructs presented in the table have been adjusted; in other words, constructs which correlated negatively with others were reversed. Various ranks are also listed: the rank of the item in the corresponding cluster (in both the $\overline{X}$ and $\overline{Z}$ grid), the rank of the item in the corresponding component of the $\overline{X}$ grid as well as of the $\overline{Z}$ grid, and the ranks of the frequencies with which the item was assigned to the dimension by cluster analysis (Q) and by principal component analysis (F) in individual grids.

The first point which deserves attention in Table 8.9 is that the various ranks do not seem to be in close agreement with each other. A closer scrutiny reveals that the ranks in cluster and the ranks in component are, by and large, of the same order of magnitude; the frequency ranks, however, look somewhat scrambled. It was pointed out earlier that low variance which was difficult to apportion or artificially high variance which could be somewhat erratic explained most of the differences between the results of cluster analysis and those of principal component analysis. Furthermore, the differences between the correlations or loadings which corresponded to two consecutive ranks were sometimes minimal; it happened that two

TABLE 8.9

SUMMARY OF THE FIRST CONSENSUS DIMENSION

ITEM Construct	POLES +	POLES −	RANK IN CLUSTER	RANK IN COMPONENT $\bar{X}$	RANK IN COMPONENT $\bar{Z}$	RANK IN FREQUENCIES Q	RANK IN FREQUENCIES F
16	Operator's fault	Nothing to do with operator	1	3	3	1	2
8	Easy to avoid consequences of danger	Impossible to avoid consequences of danger	2	1	1	4	3
14	Danger has nothing to do with design	Danger arises from bad design feature	3	2	2	2	2
2	Danger not a necessary result of process	Necessary result of process	4	4	4	3	1
17	Due to inadequate training	Nothing to do with training	1	5	5	3	3
12	Preventable	Unpreventable	5	6	6	3	4
15	Nothing to do with management	Management's fault	6	7	7	4	2
13	Anyone can put it right	Takes a specialist to put it right	7	8	8	1	3
3	Temporary danger	Permanent danger		9	9	4	2

Table 8.9 (continued)

ITEM	POLES		RANK IN CLUSTER	RANK IN COMPONENT		RANK IN FREQUENCIES	
Element	+	-		$\bar{X}$	$\bar{Z}$	Q	F
2	Donkey jacket draped across Convector heater			1	1		1
1	Main valves left open on oxy/ acetylene welder			2	2		1
17		Kegs thrown off conveyor by faulty reject arm		3	3		1
8	Gas cylinders left free-standing on ramp			5	5		2
14		Steam gushing out of tanker where people walk		4	4		4
12		Beer on floor makes it slippery		6	6		3
18		Keg lift faulty so operator has to brake hard		8	7		2
7	Tablet of soap on wet floor			7	8		4
4		Pipes across walkway		10			2
5		Badly fitting grid sticking up above floor level		9	9		5

correlations fell within the confidence limits of each other. Given these sources of error ranks in cluster and in component seem to indicate a fairly robust dimension. This casts a doubt on the reliability of frequency ranks, since the sources of potential error in frequencies are more numerous. For instance, in those individual grids where components and clusters were in reverse order, frequencies were not straightened out accordingly. Neither were the frequencies adjusted when a component was broken into two clusters. A maximum frequency of 8 meant that the figures were small, and thus a difference of 1 represented a large difference in ranks. These might be grounds for not using frequencies in subsequent analyses. However, as pointed out earlier, frequencies were useful in identifying problems which the other two methods overlooked. Therefore, it was decided that for subsequent analyses robustness would be assessed from the agreement between the two data reduction methods, and that frequencies would serve mainly as a basis for comparing consensus dimensions with corresponding dimensions in individual grids.

When Champion's (1977) results were discussed, it was argued that listing elements as well as constructs for a dimension might make the meaning of that dimension more precise. It seems to have been the case with the first dimension. What seems to differentiate the positive from the negative elements is that the positive elements seem to be the result of an operator's action or actions whereas the negative elements tend to be hazards of the environment or of the work system. It may be that the positive elements are stated as operators' faults; the differentiation may be one based on the wording of the elements. In any case, the difference between positive and negative elements corresponds to the most important construct of the dimension, i.e. "operator's fault -

nothing to do with operator".

The constructs in the first dimension appear to convey three main ideas. Firstly, whether the presence of a hazard is attributable to an operator is a major consideration. Secondly, it seems that those hazards which are deemed to be due to operators' actions are also seen by respondents as correctable by operators. Thirdly and as a consequence, these hazards are assessed as preventable.

The notion of attributability to the operator is probably the most prominent of this dimension. This is substantiated by construct 16 being most prominent on cluster analysis and by the nature and the polarity of the elements. But the elements can also be interpreted in another way.

It can be argued that positive elements are the ones within an operator's control. Hence, it would be easy to avoid the consequences of these hazards, and they would be preventable. The notions of control and of preventability seem closely intertwined; both appear to be covered by Champion's label of "scope for personal control". That the notion of "control" is an important aspect of the dimension is supported by the fact that construct 8 accounted for the highest loading in both the $\overline{X}$ and the $\overline{Z}$ grids.

Thus, it appears that the idea of danger stemming from an operator's actions is the most prominent theme of this dimension. However, the idea of control is a close second. The difference in importance between the two notions appears to be very small. But irrespective of which theme is more important, a label referring to control overlooks the first of these notions.

A label such as "scope of operator's intervention" might be more apposite. The wording may not be very elegant; but it helps to convey the two meanings of this dimension. For the substance of this dimension is about the extent to which the presence of the hazard is due to the operator's intervention. And it is also about the extent to which the operator can intervene to remove the hazard.

It is interesting to note that hazards created by the operator are deemed controllable. A parallel to this is found in Golant and Burton's (1969) second dimension. When the results of Golant and Burton's research were discussed, it was argued that the label "controllability" was not necessarily the most appropriate. Their dimension also conveyed the notion of "naturalness" of a hazard. In other words, man-made hazards were assessed as controllable and acts of God as uncontrollable.

In Golant and Burton's "controllability" dimension the two highest loadings were accounted for by scales measuring "naturalness" while the scale "controllable-uncontrollable" showed the third highest loading. So there again, "origin of the danger" seems to have been more important than "controllability". Whether or not it is the case, both notions seem intimately associated in people's mind; not only does the dimension appear with very similar contents in two independent pieces of research, but it is prominent in most individual GR grids. On average, it accounted for 27.4% of the variance of individual grids; in the $\overline{X}$ and $\overline{Z}$ grids, it explained 34.6% of the variance. So, although "controllability" did emerge as part of the main dimension, the hypothesis relating to this dimension proved to be incomplete.

8.3.2 Second dimension

"Dreadfulness" appears to be the main theme of the second consensus dimension. The specifications of this dimension are listed in Table 8.10. It can be seen from these specifications that two of the first three constructs are related to the nastiness of possible consequences of a hazard.

The presence and prominence of construct 4 ("moving - stationary danger") is somewhat puzzling. It would seem that a moving danger is seen as capable of causing more dreadful consequences than a stationary danger. This may stem from one of the few innate fears in man: the fear of something moving rapidly towards oneself (Gray, 1971). However this could also be a coincidence resulting from the elements of this dimension; all potentially dreadful dangers in this dimension were moving dangers, and no dreadful stationary danger (e.g. radiation source) was included in the list of elements of the grid. In turn, this coincidence may reflect the nature of the specific work environment from which the elements were spotted.

The elements and constructs 13, 1 and 11 also present a problem of interpretation. When these items are combined, the picture which emerges is one of rare and complex hazards as opposed to frequent and uncomplicated hazards. The former are perceived as more dreadful than the latter. It looks as though, as in the first dimension, two notions are closely associated.

It must be pointed out, however, that the last three constructs were not retained by cluster analysis in the $\overline{X}$ and $\overline{Z}$ grids. This may not be surprising for construct 13, as it was assigned to

TABLE 8.10

SUMMARY OF THE SECOND CONSENSUS DIMENSION

ITEM	POLES +	POLES -	RANK IN CLUSTER	RANK IN COMPONENT $\bar{X}$	RANK IN COMPONENT $\bar{Z}$	RANK IN FREQUENCIES Q	RANK IN FREQUENCIES F
Construct							
10	Very likely to kill	Very unlikely to kill	1	2	1	1	2
4	Moving danger	Stationary danger	2	1	2	3	3
9	Permanent disability	Only a trivial injury	1	3	3	2	3
13	Takes a specialist to put it right	Anyone can put it right		4	5	2	1
1	Never encountered in my job	Very often encountered in my job		5	4	4	5
11	Only one person at risk	Every person in the plant at risk		6	6	3	4
Element							
15	Cut-off electric eye not working			1	1		2
4		Pipes across walkway		2	2		1
11	Guard missing off machinery			3	3		2
10	Forklift truck being driven without warning light			4	4		3
12		Beer on floor makes it slippery		5	5		2

dimension 1; but the omission of the other two constructs is probably due to correlations between constructs. Whereas the first constructs seem to have some common meaning, their content is somewhat different from that of the last three constructs. On this basis, correlations between the first and the last constructs are lower than among the first constructs alone; and cluster analysis does not associate the last constructs with the first. But principal component analysis is performed on correlations between constructs and a central theme. As these correlations are likely to be higher than correlations between the constructs themselves, they may be high enough to justify inclusion of constructs which are rejected by cluster analysis.

The main theme of the second dimension seems to confirm one of the hypotheses of this research: that assessment of potential dreadfulness is the second most important consideration in assessment of hazards. This is substantiated by the nature of two of the three strongest constructs in the dimension. As described above, the other three constructs are less strongly associated with the dimension.

A detailed examination of Tables 8.9 and 8.10 reveals a curious detail. In the first dimension elements 4 and 12 are associated with one pole ("Takes a specialist to put it right") of construct 13. In the second dimension, the same two elements are associated with the other pole ("Anyone can put it right") of the same construct. It may be that these elements have in common two characteristics which are incompatible within the first dimension. For instance, it seems plausible that these two hazards are either "necessary results of the process" or not caused by the operator but can be put right by the operator. The polarity of the constructs in the first dimension would not permit the identification of such element characteristics.

Low correlations make for a less robust dimension. This second dimension explained hardly more than half the amount of variance accounted for by the first dimension; on average, the second dimension explained 19.1% of the variance in individual grids, 20.2% in the $\overline{X}$ grid and 20% in the $\overline{Z}$ grid. It must also be pointed out that again frequencies were somewhat different from the other two extraction methods. This probably indicates that dimension number 2 was less readily extracted from individual grids than was the first dimension. Frequencies of this second dimension were smaller than those of the first dimension, thus a difference of 1 represented a proportionally greater fluctuation in the frequency ranks. There are a few indications of robustness in this dimension. For instance, the three constructs retained by cluster analysis were the first three retained by principal component analysis.

It is nevertheless the case that a number of concepts are associated within this dimension. For instance, hazards which are seen as common are also seen as static and less threatening. These links between concepts deserve further attention, and will be discussed in greater depth later.

8.3.3 Third dimension

The third dimension which was extracted explained an average of 13.7% of the variance of individual grids and 13.4% of the $\overline{X}$ and $\overline{Z}$ grids. It proved to be a slightly more ambiguous dimension, as shown by its elements and constructs in Table 8.11. Poor agreement between the three extraction methods suggests that this dimension is not robust. When dimensions are compared across samples it will be argued that hazard assessment rests only on two basic dimensions,

TABLE 8.11

SUMMARY OF THE THIRD CONSENSUS DIMENSION

ITEM	POLES		RANK IN CLUSTER	RANK IN COMPONENT		RANK IN FREQUENCIES	
	+	-		$\bar{X}$	$\bar{Z}$	Q	F
Construct							
9	Permanent disability	Only a trivial injury		1	1	3	1
1	Very often encountered in my job	Never encountered in my job	1	3	3	2	1
3	Permanent danger	Temporary danger	1	4	4	1	1
5	Easily spotted danger	Very difficult to identify danger	2	2	2	4	1
Element							
11	Guard missing off machinery			1	1		2
13	Operator leaning across conveyor to operate machinery			2	2		1
18		Keg lift faulty so operator has to brake hard		3	3		3

other dimensions being secondary to the basic ones.

The main feature of this dimension seems to be either the extent to which a hazard is constantly present in a working environment, or the regular recurrence of a hazard. It would also appear that hazards which are permanent features of an environment are more likely to attract attention at one time or another, hence they would be more "easily spotted".

The presence of construct 9, however, is somewhat intriguing. It is difficult to understand why recurrent or obvious hazards should have potentially more serious consequences. It can be argued that construct 9 was assigned to this dimension by cluster analysis neither in the $\overline{X}$ nor in the $\overline{Z}$ grid; but this is probably because the construct was assigned to cluster number 2 in both grids.

It is also possible that a statistical artefact is responsible for the inclusion of construct 9 in this dimension. The dimension accounts for approximately 19% of the variance of construct 9. This percentage is broadly in keeping with that for the other constructs (18% twice and 25%) in the dimension. But as can be seen from Table 8.1, together with construct 12, construct 9 has the lowest variance across grids; 19% of a very low variance represents very little actual variance. Whereas the percentage may look significant, the straightforward explained variance may not mean much.

The general meaning of this dimension seems to be one of frequency of occurrence of a hazard. The label of "familiarity", affixed to this dimension by Champion (1977), may well be an accurate description of the main theme. However, labels such as "prominence" or "obviousness" might be more accurate; for instance, the presence of construct 9 in this dimension is more readily understandable if one considers "obviousness of danger" to be the central theme of this dimension.

Interpretation of this dimension is rendered difficult by the fact that it is less robust than the previous two dimensions; another obstacle to accurate interpretation is the small number of elements and constructs which constitute the dimension. In any case, the hypothesized notion of likelihood of occurrence appears to be at best but one connotation of the central theme. An analysis of the other samples may supply additional information about the connotations of this dimension.

8.3.4 Fourth dimension

The hypotheses of this research, as formulated in Chapter 5, encompassed only three major dimensions. It was pointed out in Chapter 4 that, if other dimensions did emerge, they would be minor ones and likely to be ambiguous.

This fourth dimension explains less variance than do other dimensions (10.2% on average in individual grids, 12.2% in the $\overline{X}$ grid and 12.1% in the $\overline{Z}$ grid); it is therefore a relatively unimportant dimension. And as described earlier in individual grids, its presence was far from obvious; so it may not be a very robust dimension. Agreement between the various ranks, which diminished with every new dimension, is reasonably good for the constructs of this dimension but not very good for the elements.

Possibly because there is some agreement between techniques about the constructs, the meaning of this dimension is somewhat less ambiguous than that of the previous dimension. As can be seen from Table 8.12, constructs 6 and 7 (the first two extracted by cluster analysis) and the first two elements of the dimension converge to suggest the notion of "salience" of hazards. What emerges is a distinction between "direct" and "indirect" hazards. In other words, there are hazards which are direct causes of accidents; all that is needed for an accident to happen is contact between a person and the hazards. As such, these hazards are more likely to result in an accident, and could be called continuous dangers.

In contrast, there are situations which are not hazards in themselves but which become hazards in the presence of other dangers. Such indirect hazards could be called contingent dangers. For example the fact that emergency doors are blacked out will not cause injuries under normal circumstances; there is danger only in case of fire.

The label "danger circumstances" chosen by Champion (1977) does not reflect this idea of "direct-indirect" hazards. Therefore, the label "immediacy of danger " has been chosen for this dimension. Such a lable has the additional advantage of conveying the idea of likelihood of occurrence which is present in the most prominent construct of this dimension.

Thus, it could be that this fourth dimension corresponds to the third hypothesized dimension. Analysis of the results of the other three samples may shed some light on this possibility.

TABLE 8.12

SUMMARY OF THE FOURTH CONSENSUS DIMENSION

ITEM	POLES +	POLES -	RANK IN CLUSTER	RANK IN COMPONENT $\bar{X}$	RANK IN COMPONENT $\bar{Z}$	RANK IN FREQUENCIES Q	RANK IN FREQUENCIES F
Construct							
7	Very unlikely to cause an accident	Very likely to cause an accident	1	1	1	1	3
11	Every person in the plant at risk	Only one person at risk	2	2	2	1	2
15	Management's fault	Nothing to do with management		3	3	2	1
6	Danger is dependent upon other things	Danger is immediately present	1	4	4	2	3
Element							
7		Tablet of soap on wet floor		1	1		3
3	Emergency doors blacked-out for film show			2	3		1
10	Forklift truck being driven without warning light			3	2		2

8.4 DISCUSSION OF THE RESULTS OF THE FIRST SAMPLE

It was shown in Chapter 7 that GR respondents differed markedly from other respondents in their use of rating scales and that approximately two thirds of all ratings in this sample were 1's and 5's. It was argued that this tendency to use the end points of the rating scale could be due to the fact that only these points of the scale were labelled. It was further argued that this tendency may have been enhanced by a level of cognitive complexity which was lower in GR respondents than in other respondents.

There were a number of problems associated with the use of extreme ratings. The first problem was construct rejections; a total of six constructs were rejected (one construct was rejected from three grids and three constructs were rejected from one grid). The second problem was that extreme ratings led either to very low or to artificially high variance. This in turn, posed problems of apportionment of variance; discrepancies resulted both within and between grids.

These discrepancies, it was argued, may have been magnified by cut-off points, in statistical analyses, which were too high. As a consequence, the various dimensions in individual grids were less than unanimous and the consensus dimensions were not as robust as desired. Nevertheless, pooling the results appears to have overcome some of these problems, since both the $\overline{X}$ and the $\overline{Z}$ grids yielded very similar results.

The hypothesis that lowering cut-off points would reduce discrepancies was tested. It was found that far fewer components were broken into two or more clusters. As a consequence, the results of the analysis of frequencies for clusters and for components turned out to be more coherent both with each other and with the two mean grids. Those discrepancies which did remain are probably attributable to the use of extreme or undiscriminating ratings.

Four significant dimensions were extracted instead of the postulated three. It has been hypothesized that the first dimension to be extracted would convey the idea of "controllability". It was found that controllability was only one of two major aspects present in this dimension. Respondents perceived that the extent to which a person was responsible for the existence of a hazard in the first place determined the extent to which the person could either control the hazard or avoid it.

As hypothesized, the second dimension centered around assessing the severity of potential consequences of hazards. Furthermore, there were also strong indications that the more mechanically complex hazards were also perceived as potentially more harmful than the comparatively simpler hazards. This dimension generates an interesting train of thoughts.

For instance, in the first dimension complex hazards were seen as originating from circumstances outside operator's actions. In the second dimension, these complex hazards were seen as more dangerous. The inference appears to be that hazards which people create around themselves are the ones whose consequences are trivial;

when potentially serious hazards exist, they are seen as not attributable to the person exposed.

Although this is just an inference, this relationship between a person's responsibility and severity of consequences emerges from evidence on tendencies in attribution of blame for accidents. Schroeder and Linder (1976) argue that when a severe accident happens in general the victims are not perceived to have been at fault; conversely, when a minor accident occurs, blame is often directed at one or more of the victims.

The third consensus dimension seemed concerned with the conspicuousness or the prominence of hazards. But some ambiguity remains about this dimension. It may well be that a salient concept, one which would measure precisely the meaning of the central theme, was missing from the grid. It has been hypothesized that this dimension would convey the idea of likelihood of occurrence. Instead of emerging in the third dimension, this concept of likelihood was extracted as a fourth dimension.

Whereas only three dimensions were predicted, a fourth one was found to be significant. The main characteristic of this dimension relates to the immediacy of hazards and to their likelihood of causing an accident. Thus the concept of conspicuousness would have prominence over the concept of likelihood. But as the third and fourth dimensions were not as robust as the other two, further evidence is required in order to consider them as basic facets of hazard assessment.

A number of other aspects of these findings deserve further consideration. These points can be discussed in a much better perspective if they are juxtaposed with findings from the other samples. Thus, the results of the other samples are analyzed in the next two chapters. The major findings from all samples are compared and discussed in greater detail in Chapter 11.

CHAPTER 9 : RESULTS FROM THE SECOND SAMPLE

The format of Chapter 9 is very similar to that of Chapter 8. Only individual and consensus grids of the second sample are discussed in this chapter. Comparisons between samples and other major findings are the subject of Chapter 11. For reading convenience, UG elements are listed in Appendix 3 and UG constructs in Appendix 4.

9.1 ANALYSIS OF INDIVIDUAL GRIDS

Each individual UG grid was processed by Ingrid 72, and principal component analysis and cluster analysis were performed according to the method described in Chapter 5. The results of these analyses are shown in Table 9.1. A number of points about these results deserve discussion before consensus dimensions are looked into.

9.1.1 Number of dimensions

The number of clusters (NQ) in each grid varied between three and five. This is very similar to what was found in GR grids. But there were important differences in the number of components (NF) extracted from each grid. The maximum number of components found significant was six - in two of the GR grids. In UG grids, there were never fewer than six components found significant.

The large number of components in UG grids may well be a statistical artefact. It is possible that there were more components simply because there were more constructs in UG grids than there were

TABLE 9.1

SUMMARY OF THE RESULTS OF CLUSTER ANALYSIS AND PRINCIPAL COMPONENT ANALYSIS FOR CONSTRUCTS IN UG GRIDS

Construct	UG01		UG02		UG03		UG04		UG05		UG06		UG07		UG08		UG09		UG10		UG11	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F
1		4-5	3	3-4	5	1-5	1	1	2	1-3		2-3 6-10	1	1-4	3	1-3 6	3	2-3		4	4	2-4 6-7
2	2	2-4	2	2	2	1-2	4	1-4	2	1	4	2-4		3-6	2	1-2	1	1-2	1	1-2		2-3 5
3	4	3-5	1	1-4	4	4-6 7	1	1-2		1-7	1	1		4		4-7 8	1	1		3-5	1	1
4	1	1-3 6	1	1	3	2-7		2-5 6	1	1-4 7	1	1	1	1		1	2	1-2	2	2	1	1
5	1	1-2 3	1	1-2	1	1-2 3	1	1-2	1	1-2	1	1	1	1	1	1	1	1-2	3	2-3	1	1-2
6	2	2	2	1-2	2	1-2	1	1	2	1-4	3	1-2 4-5	1	1	4	4-7	1	1-2 7		1-2	2	1-2 3
7	1	1	1	1	1	1-2 6	3	2	3	2-3 8		3-6	3	3		1-3 8	1	1		5	2	2-3 4
8	2	2-3	2	1-2 4	2	1-2	2	1-2	2	1-2 5	2	2-3	2	2-7	2	1-2 6	1	1	1	1-4	3	2-3
9	1	1-2	1	1	1	1-2	1	1-2		2-5	1	1	1	1	1	1		6		2-4	1	1-2
10	1	1	1	1	1	1-2	1	2	1	2	3	1-3 6	1	1	1	1	1	1-2	2	1-2 3		4-6
11	1	1	1	1	1	1-8	1	1-2 4	1	2-6	1	1-6 8	1	1	3	1-3 5		3-4	X	X	1	1-2 3
12	4	3	1	1-3	3	2-3 8		4-6	1	2-3		1-5	1	1	4	1-3 5	1	1-3	3	2-3	4	1-2 3
13	3	3-4 5	X	X	4	3-4	4	1-3	2	1-2 4-5		2-5 7	X	X		2-8	1	1	1	1-2	3	2
14	4	1-3	3	3	4	1-3	2	1-6	3	1-2 3	1	1-3	1	1	1	1-4	3	1-3 4-5	3	2-5	2	2-3

TABLE 9.1 (continued)

	UG01		UG02		UG03		UG04		UG05		UG06		UG07		UG08		UG09		UG10		UG11	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F
15	1	1	1	1	1	2-3		4-6 7	1	1-2 6	1	1	1	1-2	1	1-3 6	2	2	2	1-2	1	1-2 4
16	1	1	1	1-6	5	3-4 8	1	1-2 7	1	1-2	3	1-3		5	3	1-3 6		2-3 5		2-4 6	1	1-4
17		1-4 7	1	1	1	1-2 6	1	1	2	1	3	1-3 6-7	1	1-4 6	1	1-2	1	1-3	3	2-4	2	2-3
18	1	1	1	1	1	1-2	3	2	1	1-2	1	1	1	1-3	1	1-2	2	2	2	1-2	1	1
19	1	1	1	1-3 4	1	1-2	2	2-7	1	1-2	1	1-5		1-5	1	1-4	2	1-2 4-5	2	1-2	1	1-4
20	1	1	1	1	1	1-2	3	2-3 5	1	1-2 3	1	1	1	1-3	1	1-2	2	1-2 3-8	2	1-2 3	1	1
21	2	2	2	1-2	2	1-2 3	2	1-2	2	1-2 3	2	2-8	2	2-6	2	2	1	1-7	X	X	3	1-2 9
22	2	1-2	2	1-2	2	1-2 6	2	1-3	2	1-2	2	1-2	2	2	2	1-2	1	1	1	1-2	2	1-2 5-7
23	3	3-4 6	2	2-5	2	1-2 3	2	1-2 5-6	2	1-2	2	2-3	2	2-3	1	1-2 3	1	1-4	1	1	3	1-2
24	2	2	2	2	2	1-2	2	1-2 4	2	1-2	2	2-4	3	1-3	2	2-6	1	1	1	1	3	1-2 3
25	2	1-2	2	2	2	1-2	2	1-2	2	1-2	2	2	2	2-7	2	2-9	1	1-5	1	1	3	2-3
26		1-7		2-5	4	1-4 9	4	3-5		1-4	4	1-4 9		2-4 5-6	1	1-2 6		3-5 6		6		1-4 5
27	2	1-2 6	2	1-2	2	1-3 7	2	1	3	2-3	2	2	2	2	2	2-4 5	1	1-2	1	1	3	1-2 7
NQ	4		3		5		4		3		4		3		4		3		3		4	
NF		7		6		9		7		11		13		7		9		12		6		25

in GR grids. In fact, the number of components upon which the Bartlett test is performed is determined by the number of constructs. But in this thesis, few dimensions are expected to emerge as significant. This should leave a large number of constructs with a low latent root and thus artificially inflate the result of the Bartlett test.

There are a number of indications that a bias was present in the Bartlett test. For instance, the number of clusters in each grid was low. In addition in those four grids (UG05, UG06, UG09 and UG11) where more than ten components were found significant, Kaiser's criterion of a latent root of 1 would retain only six, seven, seven and six components respectively. Furthermore, in the same four grids the last components on which a construct was found to load significantly were F8, F10, F8 and F9 respectively.

As was the case with some GR grids, bias in the Bartlett test seems to have resulted from very high or very low variance in the UG grids. Grids UG06, UG09, and UG11 accounted for 3 of the 4 lowest variances. Grid UG05 showed the highest variance of all individual grids. But despite low variances no construct did yield unanimous ratings, and thus no construct was rejected.

9.1.2 Multiple assignments

Another noticeable feature of Table 9.1 is the fact that some constructs were assigned to three or four components. Eight respondents used twenty-seven constructs, two respondents used twenty-six constructs and one used twenty-five constructs. Out of those 293 constructs, sixty-three were assigned to three components and eleven to four components. In GR grids, no construct was assigned to more than two components. Neither between grids nor between constructs was a significant pattern found for multiple assignments.

Such assignments may simply be the result of the size of the grids. This seems to be substantiated by the fact that the three grids in which constructs were missing were the ones in which fewer multiple assignments occurred.

9.1.3 Use of the constructs

One of the differences between GR and UG grids lies in the way in which constructs were used by principal component analysis and cluster analysis. For instance, in some GR grids, some constructs were assigned to a dimension by neither analysis. Table 9.1 shows that in all UG grids each construct was assigned to at least one dimension.

This raises the topic of unanimity between cluster analysis and principal component analysis. A brief glance at Tables 8.1 and 9.1 seems to reveal greater unanimity between these techniques in UG grids than in GR grids. If constructs which were assigned to a dimension by neither analysis are included, then the techniques were unanimous for 55.1% (109) of the 198 constructs in GR grids. In comparison, 76.5% (224) of the 293 constructs in UG grids showed unanimity between techniques. On that basis, few discrepancies were expected between the three techniques of extraction of consensus dimensions.

9.1.4 Use of the elements

In GR grids, corresponding dimensions in individual grids were not always similar when their elements were compared. It was decided to check whether a similar phenomenon occurred in UG grids.

Table 9.2 is a summary of the component to which elements were assigned when individual UG grids were analyzed. It may be seen that, whereas each construct was assigned to at least one component in each individual grid, some elements were assigned to no components in some grids. Three elements (numbers 1, 13 and 16) were assigned to at least one component in all grids; one element (number 22) was assigned in only one grid.

Chi-square values were calculated to test whether non-assignments were distributed randomly between elements. The observed distribution was significantly ($p < .05$) different from a random distribution when element 22 was included; but there was no statistically significant difference between observed and random distributions when that element was excluded from the calculation. No explanation could be found for the systematic rejection of element 22, other than the fact that it is the only element which can be perceived as both the result of and the potential cause of an accident (see Appendix 3).

In grid UG11, twenty-three elements were not assigned to any components; this is nearly 50% more than in any other grid. However, no significant difference between observed and random distributions was found by a chi-square test.

TABLE 9.2

SUMMARY OF OUTCOME OF PRINCIPAL COMPONENT ANALYSIS ON ELEMENTS

ELE-MENTS	UG 01	UG 02	UG 03	UG 04	UG 05	UG 06	UG 07	UG 08	UG 09	UG 10	UG 11
1	3-5	3	5-6	1	3-5	4-5	3	4-5-6	1-3	1-2	1-2-3-4
2		1		1-2	1	1-2	2	1-3	2		
3		1		3	1-3		2	3	2	4	
4	1	1-2		1-3	2	1	2	1		1	
5	2	1	1	1	2	1	2-5	1			
6	1	1-2	1	1	1	1-2	2-4	1-2	1-5	1-2	
7	1-5			1	1-2	1	1	1		5	
8		1-2	2-4		2-4	2-3	1-3	6		1	
9	4			1-6	1				1-3	2-3	
10		1-4	2	1-2	2					1	
11	1-2	1-2			1-2-3	1-2			1-7		2-3
12	1-2-4-5	2	2-4	1-2	2-4	1-2-3		2-4	1-2	1-2	3
13	1-3	1	1-2	1-4	1-2	1	1	1	1-2-3	1-2-3	1
14		1	1-3		1	1-2		1-2	1	1	
15	1-6	1						1			
16	1	1	1-2	1-2	1-2	1	1	1-2	2	1-2	1-2-6
17	4	2		1-5				8	5	6	
18		2		5		2		2	6	1	1
19	2	2	2		2	2	2	7	1	1	1-5
20			1-2-3	1			3		1		1-2
21	2		1-2-3	2-4	3	1			1-4	3	1-2-4
22				1							
23			1		1		2	7			
24	2			7			2			1	
25		2			2					1	
26	1-2-7		1	1	1	2		1	1		
27	2	2	7	1			1	3	1	5	1
28	2			3		1	1		1	3	
29				3	1	3	1		1		1-4
30			4	2	5		6	2	2-3		
31	2	2	1-2		2	2		2	1-3	3	2-3
32	4	2		1-2	1	2-7			1		2-3-7
33		5			2	2		2			
34	1	1	1-2	1-2	1	1	1	1		2	1-2-4
35	1-2	1	2	1-2	2-3	1	1	1		2	1-2-4
36	2-3	2	2	1-2	1	2	2	2		1	
37	1	1	1	2-7	2	1		1-2		2	
38			1-6		3	4	7	1-2		2	

Eleven respondents rated thirty-eight elements each. Of those 418 elements, 3.6% (15) loaded significantly on more than two components. The distribution of multiple assignments between elements was compared to a random distribution and no statistically significant difference was found. But six of those multiple assignments were found in grid UG11 and this grid was found to be significantly ($p < .001$) different from the others. The distribution of multiple assignments between these other grids showed no difference from a random distribution.

9.1.5 An unusual grid

Grid UG11 was found to be different from other UG grids in many ways. Firstly, its mean construct variation was the lowest of all UG grids; the second lowest mean variation in a UG grid was twice the value of the mean for grid UG11. Secondly, twenty-five components were found significant; that was nearly twice the value of the second highest number of significant components of all UG grids. Thirdly, more constructs were assigned to three or more components in grid UG11 than was the case in any other UG grid. Fourthly, fewer elements were retained as significant by principal component analysis in grid UG11 than in other UG grids. But significantly more multiple assignments of elements were found in grid UG11 than in other UG grids.

There are no obvious reasons for such differences. The most plausible explanation appears to be that a language problem made the ratings in grid UG11 somewhat unreliable. Respondent UG11 seemed to experience greater difficulties in the use of the English language than other UG respondents. In addition, some unusual ratings were found in grid UG11; but there were not enough such ratings to justify rejecting the grid altogether. As the results of cluster analysis and principal component analysis for this grid were broadly in line with those of other grids it was decided not to remove grid UG11 from the sample.

9.1.6 Comparability of dimensions

As pointed out earlier, dimensions in UG grids appeared somewhat more robust than those in GR grids. It was argued in Chapter 8 that GR cut-off points were higher than UG cut-off points, and that this difference could account for fragmented (and thus less robust) GR dimensions.

It was also argued that higher cut-off points were, at least in part, responsible for difficulties in comparing dimensions between individual grids. Therefore, better comparability was expected in UG dimensions.

Labelling of dimensions, attempted unsuccessfully on GR dimensions, was undertaken on UG dimensions. This time, the exercise proved worthwhile. The results of this labelling exercise are summarized in Table 9.3.

Some dimensions were nearly impossible to label, most probably because of the absence of salient concepts; these are the dimensions for which nothing but a question mark appears in the table. Only a tentative label could be affixed to some other dimensions, usually because these dimensions encompassed few constructs and elements; these tentative labels are followed by a question mark in the table.

TABLE 9.3

SUMMARY OF LABELS FOR INDIVIDUAL UG DIMENSIONS

GRID	DIMENSION							
	1	2	3	4	5	6	7	8
UG01	Scope of intervention	Immediacy (severity)	Planning (forethought)	?	?	?	Prominence?	
UG02	Scope of intervention	Dangerous-ness	Planning (forethought)	?	Frequency of occurrence			
UG03	Scope of intervention	Dangerous-ness	Planning (forethought)	Frequency of occurrence?	?	Prominence?	?	Origin of danger?
UG04	Scope of intervention	Dangerous-ness	Familiarity	Planning (design)	Frequency of occurrence?	?	?	
UG05	Scope of intervention	Dangerous-ness	Planning (forethought & design)	Frequency of occurrence?	?			
UG06	Scope of intervention	Dangerous-ness	Origin of danger	Familiarity?	Planning (forethought)?	Prominence?	?	
UG07	Scope of intervention	Dangerous-ness	?	?	Origin of danger?	Prominence?	?	
UG08	Scope of intervention	Dangerous-ness	Origin of danger	?	Planning (forethought)?	Familiarity?	?	?
UG09	Dangerous-ness	Scope of intervention	Prominence	?	Origin of danger?	?	?	
UG10	Dangerous-ness	Scope of intervention	?	?	?	Origin of danger?		
UG11	Scope of intervention	?	Dangerous-ness	Origin of danger	Familiarity?	?	?	

Before interpreting these dimensions, some aspects of the labelling exercise deserve consideration. Firstly, as clusters and components were generally very similar, all labels were derived from components. However, in two grids (UG03 and UG04) some confusion existed between the first two components; in grid UG03 the two components had 16 constructs in common and in grid UG04 there were 10 constructs in common. Thus the corresponding dimensions in those grids were labelled from the corresponding clusters.

Secondly, there were two grids (UG04 again and UG05) in which, according to the constructs involved, the first cluster corresponded to the second component and the second cluster to the first component. Because cluster analysis is a more stringent technique, dimensions were labelled according to the corresponding clusters. Other clusters and components were very coherent (see Table 9.1), therefore other dimensions were labelled from corresponding components, especially since their elements provided useful information.

Finally, dimensions were labelled even when there were no corresponding clusters. This was done in order to search for identifiable although secondary trends or patterns. Most of the dimensions for which no corresponding cluster existed could not be labelled or else yielded only tentative labels.

Despite these slight problems, it is obvious from Table 9.3 that dimensions were far more similar between UG grids than they were between GR grids. These results made it unnecessary to search for common cores of constructs. But even if cores were not searched for some such cores (e.g. constructs 18, 19 and 20) could easily be identified in all individual grids. Therefore, the hypothesis that hazard assessment patterns could be identified which were used by most individuals received some confirmation from this labelling exercise.

9.1.7 A first interpretation of dimensions

Some of the labels listed in Table 9.3 are discussed later in this chapter in the light of consensus results. Some dimensions are also discussed further in Chapter 11, where results from all four samples are compared. However, some discussion of the labels in Table 9.3 is needed at this stage.

Table 9.3 indicates that the label "scope of intervention" was affixed to the first dimension of nine grids and to the second dimension of the other two grids. As pointed out earlier, some confusion existed in two grids between the first two clusters and components. Therefore, in order to specify the results more accurately, "scope of intervention" was the main theme of the first cluster in nine grids and of the second cluster in two grids; the same label could be affixed to the first component of seven grids and to the second component of the other four grids.

The label was chosen, much as for the corresponding GR dimension, to depict a twofold central theme. Firstly, this dimension reflects the extent to which a person in a hazardous situation is responsible for the existence of the hazard in the first place. Secondly, the dimension also conveys the idea of the extent to which a person can counter the hazard either by removing it or by avoiding it. The chosen label conveys both notions of antecedent control and prospective control described above.

The second most important label in Table 9.3 is that of "dangerousness". This dimension also has a twin central theme. Firstly, it generally conveys the idea of immediacy of hazards, of likelihood of an accident resulting from those hazards. Secondly, it also reflects an assessment of the severity of potential consequences of such accidents.

In all grids, except in UG01, constructs related to both aspects of the central theme were freely interspersed. In grid UG01 however, constructs related to immediacy tended to be more prominent members (e.g. to load higher in the corresponding component) of this dimension. That is why a different label was chosen for the second dimension in UG01.

Considering for argument's sake that the "dangerousness" label could easily have been chosen in grid UG01, this second dimension was also comparable in all UG grids. Its position, however, was less unanimous than that of the first dimension. It emerged as the first cluster in two grids, as the second cluster in eight grids and as the third cluster in one grid. In contrast, dangerousness was the main theme of the first component in four grids, of the second component in six grids and of the third component in one grid. Thus, the general tendency was for dangerousness to emerge as the second most important dimension.

It is interesting to note a slight difference between GR and UG results concerning this dimension. In GR results constructs relating to severity of consequences and those relating to immediacy (and possibly to likelihood) formed the second and third dimensions respectively. In UG results both themes are merged into one. The likeliest explanation of the difference is that the correlation between the notions of immediacy and severity is lower than GR cut-off points but higher than UG cut-off points.

Two other dimensions were (sometimes tentatively) identified in seven grids. The first one was broadly labelled as "planning"; the label "origin of danger" was used to describe the other dimension.

The notion of "planning" had two occasional corollaries. The first corollary was one of "forethought"; such a dimension conveyed the idea of an assessment of the extent to which forethought and planning (or lack of) prior to action were responsible for the presence of certain hazards. The second corollary was one of "design"; in such a dimension, the main theme centred around the extent to which poor planning, design or lay-out of an environment created hazards in that environment.

There were indications that both types of planning were considered as similar by UG respondents. For instance, in one grid both corollaries of planning are found associated within the same dimension. Furthermore, in at least two grids inadequate work techniques, usually associated with lack of forethought, are rated in the same way as bad design features.

Although also identified in seven grids, "origin of danger" had a higher mean rank of extraction than that of "planning" (4.9 compared to 3.7); thus, "origin" appears to be a somewhat less prominent dimension. This dimension reflects a contrast between dangers from human origin (e.g. because of poor planning, of lack of knowledge, etc) and dangers stemming from environmental or situational circumstances (e.g. inadequate lay-out, situations evolving too rapidly, etc).

There is a great deal of similarity between the contents of "origin of danger" and that of "planning". Both types of dimensions were found simultaneously in three grids; it could be, however, that labelling inadequacies might explain these simultaneous presences. As some "planning" and "origin" dimensions were fifth, sixth or even eighth dimensions in some grids, they generally explained little variance and thus were somewhat ambiguous.

The label "prominence" was used for a dimension in five grids. This dimension is very similar to the "obviousness" dimension identified in GR results. The main construct in this dimension was whether a hazard was easy or difficult to see. Occasional corollaries of this construct were whether a hazard was a direct or an indirect cause of potential injury, whether one came across such a hazard often or seldom, and whether a hazard was likely to cause an accident.

A dimension labelled "frequency of occurrence" and one labelled "familiarity" were both identified in four grids. Both were simultaneously present in one grid. These two dimensions had the same mean rank of extraction. They are very similar to each other and to the "prominence" dimension. As these last three dimensions were relatively unimportant trends, it is difficult to discuss them further at this stage; possibly the analysis of consensus dimensions may remove some of the ambiguities of these last three dimensions.

9.2 CONSENSUS DIMENSIONS

As for the previous sample, the frequency with which each construct was assigned to a dimension by cluster analysis and by principal component analysis was calculated. A grid of mean cell scores and a grid of mean Z-scores were also computed and analyzed. The results of these three techniques are shown in Table 9.4.

9.2.1 Agreement between mean grids

When the $\overline{X}$ and $\overline{Z}$ grids were analyzed for the first sample, they were found to yield very similar results both on principal component analysis and on cluster analysis. It was therefore expected that similar results would be found when UG grids were analyzed. However, as can be seen from Table 9.4 the results of both mean grids were not as similar for UG results as they were for GR results.

Discrepancies between mean grids were found to occur on seven constructs (numbers 12, 13, 15, 16, 18, 21 and 27). The possibility that unusual construct variance might have led to these discrepancies was looked into; no significant difference in variance was found on a Mann-Whitney U-test.

Another possible explanation for the discrepancies is ambiguity in the constructs. Two (numbers 13 and 21) of the three constructs which were rejected at least once were among the seven

TABLE 9.4

SUMMARY OF RESULTS FOR THE THREE METHODS OF EXTRACTING CONSENSUS DIMENSIONS FROM CONSTRUCTS

CONSTRUCT	DIMENSIONS 1		2		3		4		5		6		7		8		9		10		REJECTED		NONE	$\bar{X}$		$\bar{Z}$	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F		Q	F	Q	F
1	2	5	1	3	3	5	1	5	1	2		3		1						1			3	1	1	1	1
2	2	6	5	8		2	2	3		1		1											2	2	1-2	2	1-2
3	5	6		1		2	2	4		2		1		3		1							4	1	1-3-4	1	1-3
4	6	8	2	4	1	1		1		1		2		2									2	1	1	1	1
5	10	10		8	1	3																		1	1-2	1	1-3
6	3	9	5	7	1	1	1	3		1				2									1	2	1-2	2	1-2-10
7	4	5	1	4	3	5		1		1		2				2							3	1	1-2-5	1	1-5
8	2	7	8	9	1	3		2		1		1		1										2	1-2	2	1-2
9	8	8		6				1		1		1											3	1	1-2	1	1-2
10	8	8	1	5	1	2		1				2											1	1	1-2	1	1-2
11	8	8		3	1	2		2		1		2				3					1	1	1	1	1-5	1	1-5-8

TABLE 9.4 (continued)

CONSTRUCT	DIMENSIONS 1		2		3		4		5		6		7		8		9		10		REJECTED		NONE	$\overline{X}$		$\overline{Z}$	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F		Q	F	Q	F
12	4	6		4	2	8	3	1		2		1				1							2		3	2	2-3
13	2	4	1	5	2	3	2	3		3				1		1					2	2	2	2	2-3-4		4
14	3	8	2	3	4	7	2	2		2		1												1	1-3	1	1
15	8	8	2	6		2		2				3		1									1	1	1-2	3	2-3
16	5	7		4	2	4		3	1	2		3		1		1							3	1	1	3	1-2-3
17	6	9	2	4	2	3		3				3		2									1	1	1-6	1	1
18	8	9	2	6	1	1																		1	1-2	2	2-4
19	7	10	3	5		1		4		3				1									1	1	1-2	1	1-2
20	8	10	2	6	1	5				1						1								1	1-2	1	1-2-4
21	1	6	8	9	1	2						1		1		1		1			1	1		2	2	1	1
22	2	10	9	9		1				1		1		1										2	1-2	2	2-3
23	3	7	7	8	1	5		2		2		2												2	1-2-4	2	1-2
24	2	7	7	8	2	2		2				1												2	1-2	2	1-2
25	2	6	8	9	1	1				1				1				1						2	1-2	2	2-5-6
26	1	6		3		2	3	5		5		4		1				2					7	2	1-2-7	2	2
27	2	7	7	8	2	2		1		1		1		2										2	1-2		4-6-7

constructs which showed discrepancies between mean grids. Although some ambiguity may have existed in constructs 16 and 18, more explanations appear to be needed.

Another plausible explanation appears to be related to the relevance of constructs. An analysis of construct 2 (cf Table 7.7) revealed that, whereas most elements were perceived as being hazards, some elements were considered not to be hazards. For those elements which were not seen as hazards, the notions of multiplicity (construct 13), severity of injury (construct 21) and likelihood of death (construct 27) become redundant.

Finally, as pointed out earlier, a "planning" dimension proved to be the third most important one in individual grids. Both cluster analysis and principal component analysis have assigned construct 15 ("design") to dimension 3 in the $\overline{Z}$ grid. Principal component analysis has also assigned construct 12 ("planning") to the same dimension in both the $\overline{X}$ and the $\overline{Z}$ grids, and most frequently in individual grids. However, both constructs also proved to be important considerations in the assessment of "scope of intervention" (dimension 1) in individual grids. This double affinity of these two constructs may have led to ambiguities in consensus techniques.

9.2.2 Frequencies and mean grids

When the highest frequencies of assignment are compared to the outcome of the analysis of mean grids, some interesting details emerge. For instance, five of the seven constructs which showed discrepancies between mean grids also had highest frequencies in common. In other words, each one of those five constructs was assigned to a dimension most often both by principal component analysis and by cluster analysis in individual grids.

Furthermore, for these five constructs, the analysis of frequencies yielded identical results to those of the $\overline{X}$ grid. This might indicate that the $\overline{Z}$ grid is either a less robust technique or a more stringent one. As it appeared from the analysis of GR grids that the $\overline{Z}$ grid was more robust than the frequencies technique, it seems more likely that the $\overline{Z}$ grid is in fact more stringent than the other two techniques.

Of the twenty constructs for which both mean grids were unanimous for only four were the highest frequencies of Q and F not in agreement with each other. But for each of these four constructs, one (either Q or F) of the highest frequencies was in agreement with both mean grids; and when a highest frequency was in disagreement with the other indicators, the second highest frequency coincided with the results of both mean grids.

It was also noticed that some of the highest frequencies were 10's and 9's. The two highest frequencies in the GR results were 8's. This seemed to suggest that highest frequencies tended to be higher in the UG results than they were in GR results. A mean was calculated of all highest frequencies in both samples; the mean was found to be 6.9 for UG results and 5.0 for GR results. A t-test showed the difference between the means to be significant ($t = 4.512$, $df = 95$, $p < .01$).

9.2.3 Elements in consensus dimensions

Relatively strong agreement between extraction techniques, along with significantly higher frequencies, suggests that dimensions were more robust in UG grids than they were in GR grids. But before interpretation of these dimensions was undertaken, the extraction of consensus dimensions from elements was performed.

A summary of the results of the three methods of extracting consensus dimensions from elements can be found in Table 9.5. Two aspects of those results deserve further comment.

Firstly, eight elements were assigned to no dimension either in the $\overline{X}$ grid or in $\overline{Z}$ grid. Four other elements were used in only one of the two mean grids. It can be noticed that those twelve elements yielded highest frequencies of 4 or less. These frequencies were found to be significantly ($p < .001$) lower than the highest frequencies of those elements which were used in both mean grids.

Secondly, some discrepancies were found between the two mean grids. As pointed out earlier, four elements were used in only one of the two grids. In addition, one element (number 31) was assigned to dimension 2 in the $\overline{X}$ grid and to dimension 1 in the $\overline{Z}$ grid. Nine elements were assigned to more dimensions in one grid than in the other; another element (number 27) was assigned to one common and one different dimension in both grids. These elements are unlikely to be retained as important members of consensus dimensions. But should they be retained, allowance should be made for their lack of robustness.

TABLE 9.5

SUMMARY OF RESULTS FOR THE THREE METHODS OF EXTRACTING CONSENSUS DIMENSIONS FROM ELEMENTS

CONSTRUCT	DIMENSIONS 1	2	3	4	5	6	7	8	NONE	$\bar{X}$	$\bar{Z}$
1	4	2	6	3	5	2				3-4	4
2	5	4	1						4	1	1
3	2	2	3	1					4	1	1
4	6	3	1						3	1	1
5	5	3			1				3	1	1
6	9	5		1	1				1	1-2	1-2
7	6	1			2				4	1	1
8	3	4	2	2		1			4	2	2
9	3	1	2	1		1			6	1	1-2
10	3	3		1					6		
11	5	5	2				1		5	1-2	2
12	5	9	2	4	1				1	2	2
13	11	4	3	1						1-2	1-2-3
14	7	2	1						4	1-2	1-2
15	3					1			8	1	
16	10	7				1				1-2	1
17	1	1		1	2	1		1	5		
18	2	3			1	1			4	2	2
19	3	6			1		1		1	2	2-6
20	4	2	2						6	3-4	3
21	4	4	3	3					3	3	3
22	1								10		
23	2	1					1		7		1
24	1	2					1		7		
25	1	2							8	2	
26	6	2					1		4	1-2	1-2
27	4	2	1		1		1		2	2-4	2-3
28	3	1	2						5		
29	4		2	1					5		
30		3	1	1	1	1			5		
31	2	7	3						2	2	1
32	3	4	1	1			2		4	1-2-6	2
33		3			1				7	2	
34	9	4		1					1	1	1
35	7	6	1	1					1	1-2	1-2
36	3	7	1						2	1-2	2
37	5	4						1	3	1	1
38	2	2	1	1		1	1		5		

9.3 INTERPRETATION OF CONSENSUS DIMENSIONS

The same method which was used to interpret GR consensus dimensions was also used to interpret UG consensus dimensions. In other words, constructs and elements in a given dimension were ranked according to the order in which they were retained by principal component analysis and cluster analysis in the $\overline{X}$ and $\overline{Z}$ grids and in the frequencies technique. But whereas the results of cluster analysis were identical in both GR mean grids, there were some differences between clusters in UG mean grids. Therefore, the results of both cluster analyses will be used in the interpretation of UG consensus dimensions.

9.3.1 First dimension

The contents of the first dimension are listed in Table 9.6. Twenty constructs and 17 elements were retained as significant members of the dimension. On average, the dimension explained 29.4% of common variance in individual grids; it also explained 37.4% of the variance in the $\overline{X}$ grid and 32.1% in the $\overline{Z}$ grid.

There does not appear to be any major discrepancy between the results of the various techniques of extraction of consensus dimensions. However, agreement between techniques seems less obvious than it was in the analysis of GR results. On the basis of earlier analyses, a greater degree of agreement had been expected. But individual grids were relatively unanimous about the core of this dimension. In addition, the contents of the dimension proved fairly coherent and in line with previous findings.

TABLE 9.6

SUMMARY OF THE FIRST CONSENSUS DIMENSION

CONSTRUCT NUMBER	CONSTRUCT	POLES		RANK IN $\overline{X}$		RANK IN $\overline{Z}$		RANK IN FREQUENCIES	
		-	+	Q	F	Q	F	Q	F
5	This is the result of untidiness	Strongly agree	Strongly disagree	5	1	4	1	1	2
18	How easy/difficult is it to remove this?	Very easy	Very difficult	1	5			4.5	5
20	How easy/difficult is it to put this right?	Very easy	Very difficult	1	6	1	8	4.5	2
19	How easy/difficult is it to avoid this?	Very easy	Very difficult	2	10	1	5	8	2
10	It takes a specialist to put this right.	Strongly disagree	Strongly agree	3	4	3	7	4.5	9.5
15	This is due to bad design	Strongly disagree	Strongly agree	4	3			4.5	9.5
4	This hazard is only temporary	Strongly agree	Strongly disagree	7	2	6	3	9.5	9.5
9	This is the result of hurrying	Strongly agree	Strongly disagree	6	8	5	4	4.5	9.5
17	How hard/easy is it to see this?	Very easy	Very hard	11	7	7	6	9.5	5
11	This arises from something that is in use.	Strongly agree	Strongly disagree	12	13	9	13	4.5	9.5

TABLE 9.6(continued)

CONS-TRUCT NUMBER	CONSTRUCT	POLES		RANK IN $\bar{X}$		RANK IN $\bar{Z}$		RANK IN FREQUENCIES	
		-	+	Q	F	Q	F	Q	F
14	This arises from something not having been done.	Strongly agree	Strongly disagree	10	9	11	11	15	9.5
16	This is the result of someone's action(s).	Strongly agree	Strongly disagree	9	12		15	11.5	14.5
7	This is caused by a lack of knowledge.	Strongly disagree	Strongly agree	8	16	8	10	13	19.5
3	This is due to inadequate maintenance.	Strongly agree	Strongly disagree	13	15	10	12	11.5	17.5
6	In order to call this a hazard, one has to make some assumptions	Strongly disagree	Strongly agree		21		16	15	5
1	This is the consequence of another hazard	Strongly agree	Strongly disagree	14	11	12	9	18.5	19.5
2	This is a hazard	Strongly agree	Strongly disagree		14		14	18.5	17.5
8	This is sufficient on its own to cause an injury.	Strongly agree	Strongly disagree		17		17	18.5	14.5
23	How direct/indirect a cause of injury can this be?	Very direct	Very indirect		20		18	15	14.5
24	How large/small a hazard is this?	Very large	Very small		24		19	18.5	14.5

TABLE 9.6 (continued)

ELEMENT NUMBER	POLES –	POLES +	RANK IN $\bar{X}$	RANK IN $\bar{Z}$	RANK IN FREQUENCIES
16		Cooker switches hard to reach	1	2	2
34		Edge of bottom cupboard	2	1	3.5
13		Confusion over on & off position of switches	5	4	1
6	Knife sticking out of the drawer.		3	6	3.5
35		Edge of top cupboard	6	3	5.5
37		Position of cooker as regards opening of oven door	4	5	11.5
4	Soap suds on the floor		8	7	8
26	Broken bottle sticking out of the bin		7	9	8
14	Fat boiling over		9	10	5.5
2	Ball in the way		10	8	11.5
7	Tin left on the floor		14	11	8
11	Kettle lead in water		11		11.5
5	Bucket in the way		17	13	11.5
9		Flip flop shoes	12	15	14.5
3	Rucked up mat		15	12	16.5
23	Door of cupboard is open			14	16.5
15	Mixer overhanging the working surface.		18		14.5

The two poles of this dimension are virtually identical to those of the first GR consensus dimension and to those of the most prominent dimension in individual UG grids. The same label of "scope of intervention" appears to be apposite. One pole describes hazards which are part of the situation or the environment (e.g. cooker and mains switches, edges of cupboards); these are hazards which are not attributable to the person present in that environment (e.g. constructs 5, 9 and 15) and about which the person can do little if anything (e.g. constructs 10, 18 and 20).

The opposite pole describes hazards (e.g. knife, soap suds, etc.) which are the result of untidiness (construct 16) and which are not built-in features of the environment. These hazards are more easily attributable to the person concerned (e.g. constructs 5 and 9); they also tend to be under the person's control (e.g. constructs 10, 18 and 20).

There appears to be another connotation to this dimension. The last four constructs (numbers 2, 8, 23 and 24) convey the notion of direct hazards which do not depend on a chain of events before resulting in an accident. It can also be argued that construct 6 also reflects obviousness and hence immediacy. The assumption can also be made that hazards which are consequences of the other hazards (construct 1) are more likely to be direct rather than indirect causes of accidents; such an assumption, however, may be somewhat far fetched. The opposite pole is that of hazards which cause an accident only if other events take place at the same time. Built-in features of the environment are rated as indirect hazards and controllable hazards as direct ones.

The notion of direct/indirect hazards looks very much like GR consensus dimension 4. Its emergence as part of dimension 1 in UG results may be somewhat puzzling. But the constructs associated with this notion were among the last to be retained in the consensus dimension. These constructs were not retained by cluster analysis in either the $\overline{X}$ grid or the $\overline{Z}$ grid. Furthermore, the constructs were seldom assigned to this dimension in individual grids. All these indications seem to suggest that the significant loading cut-off point for principal component analysis may have been rather low.

There were other indications that UG cut-off loadings were low. For instance, as pointed out earlier in two individual grids there was a rather large number of constructs common to the first two components. The implications of low cut-off loadings are discussed later in this chapter. But one immediate implication is that the last four constructs in this first dimension seem to suggest that dimension 1 began to merge into dimension 2. Two of the last four constructs refer to the concept of injury; also all four constructs are among the first ones in dimension 2.

In GR results dimension 1 conveyed, among others, the notion of complexity of a hazard. It is therefore somewhat surprising that construct 13 (this is a multiple hazard) should be excluded from dimension 1 and that construct 24 (how large/small a hazard is this) should be the last one retained with indications that it might belong elsewhere. The explanation appears to be that, as will be discussed in relation to dimension 2, most respondents interpreted the notion of size of a hazard as being the size of the consequences which a hazard could lead to.

9.3.2 Second dimension

Table 9.7 contains a list of the eighteen constructs and seventeen elements comprising dimension 2. This dimension explained an average of 19.2% of the variance in individual grids, 28% in the $\overline{X}$ grid and 25.1% in the $\overline{Z}$ grid. When the thirty-five items comprising the dimension are analysed, some interesting - and sometimes surprising - points emerge.

Probably the most interesting aspect of this second dimension is the fact that it seems to encompass a number of concepts. The main theme of the dimension centers around the triviality/seriousness of a hazard. One pole of the dimension describes trivial hazards (e.g. elements 31, 32 and 36) which are unlikely to cause an accident or an injury (constructs 21 and 22). The other pole refers to more serious hazards (e.g. elements 6, 8 and 14) which are direct causes or sources of injury (constructs 8 and 23) and whose consequences can be severe (constructs 21 and 27).

Therefore, in addition to the postulated notion of severity of consequences, the more prominent constructs also convey the notions of likelihood of occurrence and immediacy of danger. Furthermore, construct 26 suggests that the more trivial hazards are the more common ones and that the more serious ones are less commonly encountered.

As pointed out earlier, this second UG consensus dimension appears to encompass the second ("severity of consequences") and the fourth ("immediacy") GR consensus dimensions. When individual UG dimensions were discussed, it was argued that the "dangerousness" dimension in each individual grid also conveyed the ideas of severity

TABLE 9.7

SUMMARY OF THE SECOND CONSENSUS DIMENSION

CONSTRUCT NUMBER	CONSTRUCT	POLES		RANK IN $\bar{X}$		RANK IN $\bar{Z}$		RANK IN FREQUENCIES	
		-	+	Q	F	Q	F	Q	F
22	How likely/unlikely is this to cause an accident?	Very unlikely	Very likely	3	5	4	4	1	2.5
24	How large/small a hazard is this?	Very small	Very large	1	2	2	3	6	6.5
21	How serious/trivial an injury could this cause?	Very trivial	Very serious	7	3			3	2.5
23	How direct/indirect a cause of injury can this be?	Very indirect	Very direct	5	6	1	1	6	6.5
27	How likely/unlikely is this to cause death?	Very unlikely	Very likely	1	4			6	6.5
8	This is sufficient on its own to cause an injury.	Strongly disagree	Strongly agree	6	9	1	6	3	2.5
25	How dangerous is this?	Not dangerous at all	Very dangerous	2	1	9	11	3	2.5
2	This is a hazard.	Strongly disagree	Strongly agree	4	7	3	5	8.5	6.5
6	In order to call this a hazard, one has to make some assumptions.	Strongly agree	Strongly disagree	8	8	5	7	8.5	9

TABLE 9.7 (continued)

CONSTRUCT NUMBER	CONSTRUCT	POLES −	POLES +	RANK IN $\bar{X}$ Q	RANK IN $\bar{X}$ F	RANK IN $\bar{Z}$ Q	RANK IN $\bar{Z}$ F	RANK IN FREQUENCIES Q	RANK IN FREQUENCIES F
26	One comes across this...	Very frequently	Very infrequently	9	10	6	2		18
18	How easy/difficult is it to remove this?	Very easy	Very difficult		11	7	8	12	11.5
12	This is the result of poor planning.	Strongly disagree	Strongly agree			8	9		17
20	How easy/difficult is it to put this right?	Very easy	Very difficult		12		14	12	11.5
19	How easy/difficult is it to avoid this?	Very easy	Very difficult		13		12	10	15
15	This is due to bad design.	Strongly disagree	Strongly agree		19		10	12	11.5
9	This is the result of hurrying	Strongly agree	Strongly disagree		14		15		11.5
13	This is a multiple hazard.	Strongly disagree	Strongly agree	10	15			14.5	15
10	It takes a specialist to put this right.	Strongly disagree	Strongly agree		16		13	14.5	15

TABLE 9.7 (continued)

ELEMENT NUMBER	POLES -	POLES +	RANK IN $\bar{X}$	RANK IN $\bar{Z}$	RANK IN FREQUENCIES
12		Faulty wire on mixer	1	1	1
36	Stocking of cups and saucers		2	3	3
31	Untidy work surface		4	4	3
32	Lady not wearing an apron		10	2	10
19	The fact that the drawer is open		3	14	5.5
13		Confusion over on & off position of switches	6	7	10
6		Knife sticking out of the drawer	9	8	7.5
11		Kettle lead in water	7	10	7.5
14		Fat boiling over	5	5	15.5
16		Cooker switches hard to reach	17		3
35		Edge of top cupboard	12	13	5.5
33	Tin opener left lying about		11		12.5
18	Kettle (electric) left on cooker ring		15	9	12.5
8		Jagged edge on tin lid	16	12	10
27		The man's vision being obscured by the box	8	15	15.5
26		Broken bottle sticking out of the bin	13	11	15.5
25	The fact that the bin is overfilled		14		15.5

and likelihood. Therefore, the label "dangerousness" appears appropriate to this second dimension.

As pointed out earlier, whether a hazard was large or small appears to have been rated on the basis of the size of potential consequences rather than of the size of the physical agent. Elements which load on one pole do not seem particularly larger or smaller agents than those which load on the opposite pole. But construct 24 seems closely associated with those constructs which reflect immediacy of danger and severity of consequences.

It was pointed out in the discussion of dimension 1 that the last four constructs of that dimension might be out of place and probably belonged to dimension 2. Table 9.7 substantiates this claim as these four constructs are found among the eight most important ones of the second dimension. At the same time, six of the last eight constructs (numbers 9, 10, 15, 18, 19 and 20) were also found in dimension 1.

When GR results were discussed, it was argued that those hazards which a person creates around him/herself were perceived as more trivial than faults in work systems or in the environment. The last constructs of UG dimensions 1 and 2 suggest a strong correlation between the two dimensions; and that correlation appears to substantiate the postulated link between human errors and triviality of hazards.

The last constructs of dimensions 1 and 2 (except number 13 in the latter) have been retained as members of the corresponding clusters neither in the $\overline{X}$ grid nor in the $\overline{Z}$ grid. It was mentioned in Chapter 6 that cluster analysis was a more stringent technique than principal component analysis. Therefore, the minimum loadings significant at the .05 level were probably lax. As a matter of curiosity, principal component analysis was performed again on UG mean grids but with the cut-off loadings used for GR grids; the new components were noticeably closer to the corresponding clusters.

As with the first consensus dimension, despite relative unanimity about the core of the second dimension, some discrepancies were observed between the three extraction methods. Although three more consensus dimensions were identified, this second dimension appears to encompass GR consensus dimensions 2 and 4. This may be due to some unreliability in the dimension. On the other hand, it may be that for some unknown reason, the demarcation line between dimensions was not stringent enough. Results from the other two samples may shed some light on this point.

9.3.3 <u>Third dimension</u>

The third dimension which was extracted from UG results is more ambiguous than the first two dimensions. That is probably because it generally explained little of the common variance: 8.5% on average in individual grids, 7.4% in the $\overline{X}$ grid and 8.7% in the $\overline{Z}$ grid. As can be seen in Table 9.8, only six constructs and 3 elements were retained as significant members of this dimension.

TABLE 9.8

SUMMARY OF THE THIRD CONSENSUS DIMENSION

CONSTRUCT NUMBER	CONSTRUCT	POLES		RANK IN $\bar{X}$		RANK IN $\bar{Z}$		RANK IN FREQUENCIES	
		-	+	Q	F	Q	F	Q	F
12	This is the result of poor planning	Strongly disagree	Strongly agree		1		4	3	1
16	This is the result of some-one's action(s)	Strongly disagree	Strongly agree			1	2	3	3
14	This arises from something not having been done	Strongly disagree	Strongly agree		4			1	2
15	This is due to bad design	Strongly disagree	Strongly agree			1	1		5.5
13	This is a multiple hazard	Strongly disagree	Strongly agree		3			3	4
3	This is due to inadequate maintenance	Strongly disagree	Strongly agree		2		3		5.5

ELEMENT NUMBER	POLES		RANK IN $\bar{X}$	RANK IN $\bar{Z}$	RANK IN FREQUENCIES
	-	+			
1		Too many things going on at once	1		1
21	The way the lady is holding the vegetable for cutting		2	1	2
20	Lady not looking at what she is doing		3	2	3

The main theme of the dimension appears to be contained in the most important construct (number 12) of the dimension. The dimension as a whole seems to refer to an assessment of a global situation as opposed to various activities within it. One pole of the dimension appears related to a system of work whereas the other pole seems concerned with particular activities within the system.

A label such as "lack of systems" might be appropriate, thus encompassing construct 3 if maintenance is seen as a preventive measure. However, it could well be that the main theme of this dimension is the same as that of the third most important dimension in most individual UG grids. The label "planning" would appear very apposite. The dimension also encompasses the two main corollaries of planning ("forethought" and "design") which were discussed earlier.

Although interesting, this dimension remains slightly unreliable. For instance, there was no third cluster in the $\overline{X}$ grid. And there were some important differences between the other techniques about the contents of the third dimension. However, a very similar dimension was present in seven individual grids. This in itself lends some credibility to this dimension. Thus it could be that the hypothesis relating to the third dimension might prove to be wrong. Evidence from the other two samples might shed some light on this point.

9.3.4 Fourth and fifth dimensions

Two more dimensions were extracted, but with some difficulty. Their items are listed in Tables 9.9 and 9.10 respectively.

TABLE 9.9

SUMMARY OF THE FOURTH CONSENSUS DIMENSION

CONSTRUCT NUMBER	CONSTRUCT	POLES -	POLES +	RANK IN $\bar{X}$ Q	RANK IN $\bar{X}$ F	RANK IN $\bar{Z}$ Q	RANK IN $\bar{Z}$ F	RANK IN FREQUENCIES Q	RANK IN FREQUENCIES F
13	This is a multiple hazard	Strongly disagree	Strongly agree		1		1	1.5	1
3	This is due to inadequate maintenance	Strongly agree	Strongly disagree		2			1.5	2

ELEMENT NUMBER	POLES -	POLES +	RANK IN $\bar{X}$	RANK IN $\bar{Z}$	RANK IN FREQUENCIES
1		Too many things going on at once.	1	1	

TABLE 9.10

SUMMARY OF THE FIFTH CONSENSUS DIMENSION

CONSTRUCT NUMBER	CONSTRUCT	POLES		RANK IN $\bar{X}$		RANK IN $\bar{Z}$		RANK IN FREQUENCIES	
		-	+	Q	F	Q	F	Q	F
11	This arises from something that is in use	Strongly agree	Strongly disagree		1		1		1.5
7	This is caused by a lack of knowledge	Strongly agree	Strongly disagree		2		3		1.5

The fourth dimension encompassed only two constructs and one element. It explained an average of 6.7% of the variance in individual grids, 5.5% of the variance in the $\overline{X}$ grid and 6.6% in the $\overline{Z}$ grid. The main concept in the dimension appears to be the complexity of hazards; this is outlined by construct 13 and element 1. It is difficult, however, to explain why construct 3 is associated with the other two items of the dimension.

The fifth and last dimension is even more obscure. It only explained an average 5.6% of the variance of individual grids, and 4.2% and 5% respectively in the mean grids. One pole of the dimension appears to convey the notion that something is being done in the wrong way. But little more can be said about this dimension.

The evidence available on these two dimensions is insufficient to clarify their possible meanings. Their contents rarely match those of the minor dimensions tentatively identified in individual grids. The small amount of variance explained and the low frequencies suggest that these dimensions are rather unreliable. Since the corresponding individual dimensions were also weak, the possibility that fourth and fifth dimensions may be statistical artefacts can certainly not be excluded. As pointed out in Chapter 11, only the first three dimensions find their equivalents in other samples.

9.4 DISCUSSION OF THE RESULTS

Consensus dimensions identified from UG results are discussed in greater detail in Chapter 11 when they are compared to dimensions from other samples. However, a brief discussion of the main points in this chapter may outline topics on which the results of the last two samples may shed some light.

UG respondents rated a grid which was approximately three times the size of the grid rated by GR respondents. Yet, all constructs in each individual UG grid loaded significantly on at least one component. Two explanations can be suggested for such a difference in the use of constructs.

Firstly, as UG grids were larger than GR grids, the minimum loading to be significant at the .05 level was lower for UG than for GR components. A lower minimum loading may have meant that, assuming identical loadings in UG and GR components more constructs could be retained as significant members of UG components.

This argument finds support from the fact that UG consensus dimensions were less clearly defined than GR dimensions; it was argued that such imprecisions were due to minimum loadings which were too low. For example, a minimum loading of 0.325 for component 1 meant that a UG construct was retained when construct and component only had 10.6% of common variance. Nearly twice that amount of common variance was needed for a GR construct to be retained as a member of the same component.

It was argued in Chapter 8 that GR cut-off points may have been too high. Combining high GR loadings and low UG cut-off points probably explained why two GR consensus dimensions were merged into one UG consensus dimension. But lower minimum loadings in UG grids do not explain all the differences between the results of the two samples.

An additional explanation is probably that UG respondents made a fuller use of the rating scale than did GR respondents. It was pointed out in Chapter 7 that GR respondents tended to use rating scales dichotomously whereas UG respondents used intermediate ratings (2's and 4's) more. It was argued in Chapter 7 that a greater level of cognitive complexity enhanced the effect of labelling all points of each scale, which led to such different rating patterns. A further possibility is that UG respondents, having themselves elicited the constructs they used, understood the finer points of those constructs better.

There were far fewer constructs unassigned to any dimension in UG grids than there were in GR grids. Again this could be because cut-off points were too high in GR grids or too low in UG grids. Unassigned constructs could be "latent" dimensions, i.e. salient concepts of dimensions which do not encompass enough constructs to emerge as significant. But if unassigned constructs were latent dimensions they would probably have generated smaller, only just significant dimensions. It appears more likely that unassigned constructs are ideas which do not belong to people's main line of thinking on hazards.

The number of components extracted from individual UG grids was greater than the number of components extracted from GR grids. But there are indications that too many components were extracted from UG grids. For instance, the number of clusters in UG grids was comparable to that in GR grids. Moreover, only one more consensus dimension was extracted from UG data than was extracted from GR data; and this last UG consensus dimension was hard to define accurately.

The core of the first UG consensus dimension closely resembled the first GR dimension. In this dimension, two main ideas were associated: the person's role in the occurrence of a hazard and the possibility for the person to keep a hazard under control. There were indications that to a certain extent the first dimension merged into the second dimension; the notion of immediacy of danger, very prominent in the second dimension, was also present in the last constructs of the first dimension.

In addition to the notions of immediacy and likelihood of occurrence, concepts related to severity of consequences were also present in the second dimension. Constructs dealing with environmental faults as opposed to human error were also found in the second dimension; but it seems likely that their presence in that dimension was an artefact. Therefore, it looks as though the concept of dangerousness was the key one in the second dimension, although construct 25 had only the seventh highest loading in the dimension.

Again in UG results the concepts of controllability and severity of consequences were found in the first and second dimensions respectively. But, as in GR results, other concepts were associated with them. For instance, in both series of results, environmental faults were perceived to be both uncontrollable and unpreventable. A further similarity between UG and GR results is that what is seen as uncontrollable is also seen as harmful. In other words, there is a strong correlation between dimension 1 and dimension 2.

There were also differences between the results of the two samples. For example, to a certain extent UG dimension 2 appears to encompass GR dimensions 2 and 4. This may be because there is such a strong correlation between the two GR dimensions that a rather lax demarcation line in UG analyses failed to discriminate between them.

Another dimension emerged from UG data which had not been postulated and which had not been identified in GR grids. This dimension centered around an assessment of the extent to which a hazard is related to planning or to a fault in planning - whether planning one's actions or planning the design of an environment. This dimension, although somewhat lacking in robustness, can lead to interesting speculation. This is discussed at greater length in Chapter 11, where dimensions extracted from the four sets of grids are compared.

Two more dimensions were identified in UG results. Neither appears related to the dimensions which were postulated. Furthermore, neither seems comparable to any of the GR consensus dimensions. But the contents of these two dimensions were so limited, and they explained so little common variance in UG grids, that their main themes were almost impossible to grasp.

CHAPTER 10 : RESULTS FROM THE THIRD AND FOURTH SAMPLES

The first two samples (UG and GR) were the only ones for which the full range of systematic analyses were possible. Because SA respondents rated common elements on different constructs, no consensus grid could be calculated. As for the fourth (PH) sample, a very small grid (six constructs and nine elements) was rated by three respondents only. Nevertheless both samples yielded very important results.

It was decided to review the results of the last two samples within one chapter. This chapter has two main sections, one devoted to the results of each of the remaining samples. The outlines of each section will be similar to those of Chapters 8 and 9. Extended comparisons between samples will be the subject of the next chapter.

10.1 RESULTS FROM THE THIRD SAMPLE

As described in Chapters 4 and 6, SA respondents (17 students at the Dundee School of Architecture) rated their grid individually. In other words, each respondent was interviewed separately. During the interview, constructs were elicited from a standard list of hazards used as elements (see Table 4.3 and Appendix 5). Therefore no two grids included exactly the same constructs. There were twenty-one elements, and the number of constructs varied between fifteen and twenty-five.

10.1.1 Processing of individual grids

Individual grids were processed as usual by Ingrid 72. Principal component analysis and cluster analysis were performed for each grid.

A number of problems arose even at this early stage. Firstly, eleven of the seventeen SA respondents used ratings of zero to signify that some constructs were not applicable to some elements. It was found out that Ingrid 72 treated ratings of zero as the lower extreme of the rating scale instead of excluding such ratings from further analyses. Before rejecting grids with large proportions of zeros it was decided to test whether ratings of zero had an effect on the extraction of underlying dimensions. Grids were sorted into those with no, those with few and those with many ratings of zero, and the results of the three groups were compared. These comparisons are described later in this chapter.

Secondly, it was found that in two grids the Bartlett test of the number of significant components was not performed. The possible causes of these omissions were sought. What emerged was that in a number of cases, two elements were rated identically within a grid, thus causing correlations of 1.0 between these elements. There were also constructs in some grids between which correlations of 1.0 were found.

Four different pairs of elements (in four different grids) were found to have been rated as identical. In a first grid "train crash" was perceived as identical to "plane crash". In a second grid, "plane crash" received the same ratings as "accident in a chemical plant". This last hazard was perceived as the equivalent of "accident on a building site" in a third grid. Finally, in a fourth grid, "rock climbing accident" and "skiing accident" had a correlation of 1.0.

One pair of constructs with a correlation of 1.0 was found in three grids. In another grid, three pairs and a group of three constructs were found to measure exactly the same thing. Finally, three pairs, a group of four and a group of five constructs with maximum correlations were found in another grid.

It was thought necessary to retain only one element or one construct from each pair or group of identically rated items; all other items were removed from the corresponding grids. Finally, constructs on which all elements received identical ratings were also removed. The grids affected by these removals were re-processed.

Once all the grids were correctly processed, the next step was to interpret the results. As pointed out earlier, because no two grids comprised identical series of constructs, it was impossible to calculate a consensus grid. Some methods could nevertheless be used to draw some conclusions from the sample as a whole, although these methods were less rigorous than a consensus grid. They were: a broad comparison of dimensions extracted from individual grids and an analysis of those constructs which could be found in more than one grid. These analyses are described in forthcoming sections of this chapter.

All these analyses necessitated a reference, in one form or another, to the nature of the dimensions extracted from individual grids. It was therefore necessary to test the robustness of these dimensions prior to the analyses of the sample as a whole.

10.1.2 Agreement between clusters and components

When the results of cluster and principal component analyses were compared, a high level of agreement between the two techniques was found. Two methods were used to test this agreement.

Firstly, correlations were calculated between the number of clusters and the number of components extracted from each grid. A parametric correlation of .839 and a non-parametric rank correlation of .819 were obtained; both correlations were found to be significant well beyond the .001 level. Therefore, the number of clusters in each grid was closely proportional to the number of components.

The next step was to check whether the contents of the clusters resembled the contents of the corresponding components. In three of the grids, there were only one cluster and one component, very similar to each other in all three cases. In seven other grids, the first cluster and the first component comprised practically the same constructs. In five more grids, the first component was the amalgamation of the constructs of the first cluster and of another cluster. Such a fragmentation of a component into more than one cluster is not surprising considering that cluster analysis is the more stringent technique of the two. Finally, in the last two grids, the first cluster was also extracted as a component, but not as the first one; and conversely, the first dimension was also extracted as a cluster, but not as the first one. In these two grids, small first latent roots (one of them being the lowest of the sample) probably explain these inversions of dimensions.

There were, therefore, very few discrepancies between the results of the two types of analyses. Thus, the dimensions from SA grids discussed in subsequent sections of this chapter can be considered relatively robust. In the few cases where there were discrepancies, the nature of the dimensions was interpreted from the contents of the components. The components comprised elements as well as constructs, which facilitated their interpretation.

10.1.3 Comparisons of individual dimensions

Table 10.1 presents a listing of labels affixed to the various dimensions in individual grids. The number of clusters (NQ) and the number of components (NF) are also provided.

A number of technical comments about the contents of Table 10.1 need to be made before dimensions are compared. Firstly, 16 dimensions (identified by question marks) could not be labelled. Seven of these dimensions were found to be significant according to the Bartlett test but no construct and no element loaded significantly on them. The other nine dimensions were not labelled because their contents were either insufficient or very ambiguous. Secondly, in six grids the number of labelled dimensions was closer or equal to the number of clusters; in three grids, the number of labelled dimensions

TABLE 10.1

SUMMARY OF THE LABELS AFFIXED TO DIMENSIONS IDENTIFIED IN INDIVIDUAL SA GRIDS

DIMENSION NUMBER	SA01	SA02	SA03	SA04	SA05	SA06	SA07	SA08	SA09	SA10	SA11	SA12	SA13	SA14	SA15	SA16	SA17
1	Hideousness	Severity of consequences	Riskiness	Severity of consequences	Widespread/ localised	Widespread /localised	Severity of consequences	Horror	Severity of consequences	Likely/ unlikely	Controll- ability	Enclosed/ open-air	Foresee- ability	Number affected	Number affected	Severity of consequences	Likely/ unlikely
2		Natural/ man-made	Controll- ability		Preventab- ility	Enclosed/ open-air	Number affected		Vol ntar- iness of activity	Natural/ man-made	Widespread/ localised	Natural/ man-made	Voluntar- iness of activity	Severity of con- sequences	Severity of conse- quences	Natural/ man-made	Voluntar- iness of activity
3		Occupation- al/not occu- pational	Number affected		Controll- ability	Natural/ man-made	Modern/ traditional danger		Natural/ man-made	Controll- ability	Number affected	Easy/ difficult escape	Prevent- ability	Short-term /long-term effects	Widespread/ localised	?	?
4		Widespread/ localised	Acceptance of risk		Voluntari- ness of activity	?	Likely/ unlikely		Enclosed/ open-air	Voluntar- iness of activity	?	Voluntar- iness of activity	Controll- ability	Familiar- ity	Voluntar- iness of activity	Extent of material damage	
5		Ugliness			?	?	Size of pop- ulation at risk		Moving/ stationary	Riskiness	?	Moving/ Stationary	Attribut- ability of blame	?	Acceptance of risk	Preventab- ility	
6					?		Voluntar- iness of activity		Modern/ traditional danger	?		Familiarity	?		Nature of consequences	?	
7									?				?		?	?	
NQ / NF	1 / 1	4 / 5	2 / 4	1 / 1	4 / 6	3 / 5	6 / 6	1 / 1	6 / 7	4 / 6	3 / 5	5 / 6	4 / 7	3 / 5	4 / 7	3 / 7	2 / 3

was half-way between the number of clusters and the number of components; in four grids, the number of labelled dimensions was closer or equal to the number of components. The unlabelled dimensions were all found in grids where there were more components than clusters, and they explain 16 of the 25 components not matched by corresponding clusters. Thus the vast majority (57 out of 66) of labelled dimensions corresponded to both a cluster and a component, and as such they can be considered as robust.

Labels were generally chosen on the basis of the most prominent construct in the dimension. For eleven of the dimensions, however, this was not the case. In all eleven cases, however, the main notion of the dimension corresponded to a construct which was among the first four members of the component and the first three members of the corresponding cluster, and which seemed to be the best way of describing the difference between the positive and negative elements.

Twenty-three labels were used for the 66 dimensions which could be interpreted. One label was used for eight dimensions, one for seven dimensions, one for six dimensions, three labels were used five times, four labels were used three times, five were used twice and eight were affixed only once.

The most frequently used label was "Voluntariness of activity". In such a dimension hazards are assessed as a function of whether the individual has the choice of engaging in the activity during which danger arises. Although this was the most frequent dimension, it tended to be a relatively unimportant one. In the eight grids where it is found, it comes in second position three times, in fourth position four times and in sixth position once.

The second most frequently affixed label was found in seven grids. The dimension generally assesses the "severity of consequences" which can result from the various hazards. This severity is estimated in terms of bodily damage, lives lost, material damage and/or financial cost. In three of the seven grids, there was an additional, less prominent dimension assessing specific aspects of the severity of consequences. Thus, in grid SA14, there is a dimension which assesses whether the consequences or effects spread over a short or a long period of time. In grid SA15, the "Nature of consequences" dimension is concerned with whether the consequences are mostly bodily damage or material damage. The dimension labelled "Extent of material damage", in grid SA16, assesses only the amount of material damage which can result from the various hazards; this dimension is not concerned with bodily damage or injuries. In other grids, there is a dimension, labelled "Number affected", which looks at another specific aspect of "Severity of consequences"; but as "Number affected" is one of the frequently used labels, this dimension is discussed in detail later. In those grids from which a "severity of consequences" dimension was extracted, this dimension tended to be very prominent; in five grids, it was extracted as the first dimension (in grid SA04, it was the only dimension identified) and it was the second dimension in the other two grids.

The third most frequent label was "Natural/man made"; it was affixed to a dimension in six grids. This dimension examines the origin of danger and whether some human intervention had a role in this origin. This dimension appears to be somewhat less prominent than "severity of consequences", but a little more prominent than "voluntariness of activity". It was extracted in second position from four grids and in third position from two other grids.

There were three dimensions which were identified five times each. It is interesting to note that one of these three (i.e. "controllability") is relatively close to the preceeding ("natural/man-made") dimension. The similarity between these two dimensions is discussed later on in this chapter. The "controllability" dimension is concerned with the extent to which human intervention can prevent the occurrence of danger or avoid the consequences if a mishap occurs. The prominence of extraction of the "controllability" dimension was about the same as that of the "natural/man-made" dimension. The "controllability" dimension was extracted in first place once, in second place once, in third place twice and in fourth place once.

The second dimension to be extracted five times was given the label "number affected". This dimension is concerned with the number of persons killed, injured or suffering a financial loss as a consequence of a disaster. The label "kill size" was also considered, but it was found to be somewhat too narrow to reflect the full meaning of this dimension. For instance, the dimensions labelled "number affected" referred to "effects" rather than to "lives lost", and to "number of people affected" and to "injury results" rather than to "number killed". Thus, the label "number affected" appeared more apposite. This dimension was found in first position twice, in second position once and in third rank twice.

"Widespread/localised" was the label chosen for the third dimension which was found in five different grids. This dimension attracted at one pole hazards which were diffuse or disseminated (e.g. "air pollution" and "nuclear radiation") or which affected a geographical area (e.g. "earthquake"); at the other pole were found hazards which occurred in a precise location, in a specific spot (e.g. "rock-climbing accident", "car crash"). This dimension was extracted in first position twice and in second, third and fourth rank once each.

Three dimensions were each found in three grids; their labels are "likely/unlikely" (first rank twice and fourth rank once), "preventability" (second, third and fifth ranks) and "Enclosed/open-air" (first, second and fourth positions). Their prominence is similar to that of other dimensions, but the fact thay they are found in only three grids each suggests that they may be, on the whole, somewhat secondary dimensions. They are nevertheless discussed further in subsequent sections of this chapter.

Five more dimensions were found in only two grids each, and eight dimensions were each extracted from only one grid. Not only were these dimensions infrequent, but they also tended to be amongst the last to be extracted from their respective grids. Thus, it seems very likely that they are very minor dimensions. There are, however, three dimensions which, even though found only once each, deserve further comments. All three reflect a certain dread inspired by hazards. The labels "hideousness", "ugliness" and "horror" are more specific than "dreadfulness", but the notion of dread appears to be the common denominator of all three dimensions. It is also interesting to note that in two grids (SA01 and SA08) the dimension with an undertone of dread was the only one to be extracted, explaining over 75% of the variance in both cases. This point is the subject of a lengthier discussion in the section of this chapter where these results are interpreted.

10.1.4 Comparisons of elements

There is usually a certain amount of subjectivity (and hence a possibility of bias) in labelling dimensions. Thus other means were sought in order to remove some of the subjectivity involved in the labels discussed above and to substantiate the attempt at identifying hazard assessing patterns emerging from the SA sample in general. Firstly, it was decided to investigate the behaviour of the elements in individual grids; this investigation is the subject of the present section. Secondly, it was thought useful to look at the behaviour of comparable constructs across individual grids; this part of the study is discussed in the next section of this chapter.

There were two ways in which the behaviour of elements could be studied. Firstly, those pairs and triads of elements which could be found most often together, in one dimension across grids, were identified. Larger groupings were looked for but very few could be found in more than two grids; those which were found were related to the pairs and triads already identified. All related groupings of elements are discussed together. Secondly, similarly labelled dimensions were compared to find out whether their elements were similar. These two approaches are now discussed in turn.

One of the important considerations in identifying frequent triads was to make sure that the elements of each were at comparable poles. In other words, if in one grid two elements were found at the positive end of a dimension and a third element was found at the negative end, the triad was considered to be identical to that in the next grid if the two previously positive elements were again positive and the negative element remained negative. The triad could also be considered valid if the two positive elements in one grid had become negative, provided that the previously negative element had also changed sign.

Bearing these conditions in mind, the most frequently identified triad involved elements number 7 ("accidental release of nuclear radiation") and 10 ("earthquake") at one pole and number 16 ("skiing accident") at the opposite pole. This triad was found in eight individual grids. A search was carried out in order to determine which other elements had a tendency to be extracted along with the previous three. Two elements, i.e. numbers 12 ("rock-climbing accident") and 21 ("being knocked down while crossing the road") were often found at the same pole as element 16.

There were ten possible combinations of three elements among the five mentioned above. As pointed out earlier, one of these combiantions was found in eight grids, five combiantions were found in seven grids, three combinations were found in six grids and one combination in five grids. Thus it is obvious that more than one combination could be found within some grids; this meant that many grids united within one dimension more than three of these five elements. So, combinations of four elements were looked into. There were five such possible combinations; three of them were found in six grids and two were found in five grids. The full set of five elements was extracted from five individual grids. Some of these figures (including the last one) can be increased by one, considering that in grid SA15, elements 12 and 16 (rock-climbing and skiing accidents) *received identical ratings* and displayed a perfect correlation of 1.0.

Before undertaking a search for another group of frequently associated elements, a comparison was made of the labels of the dimensions in which the five elements were found. There were, in all, nine grids in which these elements were found in groups of four or five. In four of these grids, the elements were found in the dimension labelled "severity of consequences"; in two grids, the label of their dimension was "controllability"; in the remaining three grids, the labels "widespread/localised", "horror" and "voluntariness of activity" were used once each to describe the dimension encompassing the elements. Elements 7 ("accidental release of nuclear radiation") and 10 ("earthquake") were assessed as having the most severe consequences, as being outside one's control, horrific and widespread, and as not necessarily occurring during an activity voluntarily engaged in. Conversely, the other three elements were associated with the other poles of these dimensions. Further comments about these associations are made later in this chapter.

The search for other groups of elements led to the identification of a second frequent triad. It comprised the elements number 8 ("home fire") and 18 ("accident in the home"), associated at one pole, and element number 9 ("plane crash") at the opposite pole. This triad was found in six grids. Two elements were found to join this triad in four grids; they were numbers 5 ("illness") and 19 ("food poisoning"), both at the same pole as elements 8 and 18. All five elements were found together in one dimension in three grids.

Strong associations were found between this group of elements and the first cluster of elements discussed above. There were three types of indications that such associations existed. Firstly, the second cluster of elements was often associated with some elements (in particular numbers 7, 10 and 16) of the first group. Secondly, in three of the six grids the elements were extracted as members of the same dimension as that to which elements of the first group belonged. Thirdly, as argued below, the labels of the other three dimensions were very similar to the labels associated with the first group of elements.

There were six grids in which at least four of the elements of the second group were clustered. As pointed out above, three of these dimensions have already been mentioned in relation to the first group of elements. Their labels were: "controllability", "severity of consequences" and "horror". As for the other three dimensions encompassing the elements of the second cluster, their labels were "hideousness", "controllability" and "widespread/ localised". These last two labels have also been used for elements of the first group, whereas "hideousness" is somewhat akin to the "horror" dimension encompassing both groups of elements.

No other group of elements stood out quite as clearly as the previous two clusters. There were strings of up to eleven elements which could be found in more than one grid. But the majority of these clusters were made up of the ten elements already mentioned; the longer strings (eight elements or more) were found in only two grids. Other minor groups of two, three of four elements were identified. However none of these minor groups were found in four or more grids; furthermore, they tended to belong to larger strings of elements encompassing a majority of elements among those of the two main clusters.

Some reservations must be expressed at this stage about the method described above. In contrasting and comparing elements the aim was to find a characteristic, and thus a construct, which explained their arrangement as it was identified. The risk in such a procedure was that the researcher applied his own construction system to explain the arrangement of elements. Furthermore Kelly (1955) argues that it is not so much the objective elements as the way they are construed which is important in understanding a person's perception or assessment of the world. Thus, identical objects can be construed differently by different people. The fact that identical elements were found in differently labelled dimensions may well be an indication that these elements were not construed in the same way by all respondents. It seems that greater emphasis should be placed on the interpretation of constructs. Nevertheless, not pursuing the analysis of elements might bead to overlooking useful information; in the first two samples, element contents were occasionally needed to interpret dimensions which constructs alone could not clarify. Therefore, it was decided that attention should be focused upon construct contents, but also that the analysis of elements should serve as complementary information when needed.

As pointed out earlier, identical groups of elements were found in differently labelled dimensions. This raised the question of whether the contents of dimensions with identical labels were comparable. Therefore, the second type of analysis of the elements was carried out: the contents of similarly labelled dimensions were compared to see whetehr they had some elements in common.

During these comparisons contradictory contents were looked for. Special attention was paid to elements which might load on one pole in one dimension and on the other pole in the similarly labelled dimension of another grid. Of the 213 significant element loadings in the various dimensions found more than once, only 15 loadings were found to be in contradiction with loadings of the same elements in comparable dimensions. These could indicate either that similarly labelled dimensions were not really comparable, or that the specific elements were construed differently.

Raw data (original ratings) were consulted to check the latter possibility. It was found that most contradictory ratings stemmed from different ratings, and thus from dissimilar construing. For instance "coal mining accident" was rated by one respondent as killing few people, and by two respondents as killing many people. This suggests that the concept of "many/few killed" does not have the same meaning for everyone. Another example is "fire in a discothéque" rated by one respondent as voluntary and by three respondents as involuntary. This difference in construing raises an interesting point. It seems likely that what one respondent rated as voluntary was the activity during which a hazard occurred (i.e. being in a discotheque) whereas other respondents rated the hazard itself (i.e. the fire) as being encountered involuntarily. This point is discussed again later in this section.

Other inverted loadings were found to reflect the exact opposite of the relevant raw data. It was found that these inversions occurred for elements which received a few ratings of zero, meaning that the constructs eliciting ratings of zero were not considered relevant to the assessment of the elements. Ratings of zero caused other problems which are discussed later in this chapter.

Comparisons of elements contents were, on the whole, disappointing. For instance, there were eight grids in which a dimension was labelled "voluntariness of activity"; no element was found to be a member of this dimension in more than three grids. "Home fire" and "Hotel fire" were both found in three grids as examples of events which occur while a leisure activity is being carried out; in contrast, in three grids "illness" was considered as an event which could occur irrespective of the type of activity being carried out. These three elements were not necessarily found in the same three grids. As pointed out earlier it seems possible that the concept of voluntariness may have been interpreted differently by various respondents. For instance, it may be that some respondents assessed the voluntariness of the activity itself whereas other respondents assessed the voluntariness of the encounter with the hazards, i.e. the extent to which the hazardousness of an activity is known before the activity is undertaken. Possibly some respondents sorted the elements into those activities which they would and those activities which they would not voluntarily partake in themselves, whereas other respondents assessed the extent to which people generally choose to undertake an activity on a voluntary basis. An examination of the construct contents is needed to shed some light on this point.

Those dimensions which were labelled "severity of consequences" (in seven grids) turned out to have contents more similar than those of the previous dimension. "Earthquake" in six grids and "accidental release of nuclear radiation" in four grids were considered as leading to severe consequences; "skiing accident" in five grids, "being knocked down while crossing the road", "accident on a building site" and "accident in the home" (each found in four grids) all represented events leading to less severe consequences. Six other elements were found in three grids each.

The contents comparison of the six "natural/man-made" dimensions was not quite as fruitful. "Being struck by lightning" in four grids and "earthquake" in three grids were the only frequent examples of natural disasters. As for man-made hazards, examples varied from one grid to another.

The label "widespread/localised" was affixed to a dimension in five grids. In four of these grids "air pollution" and "accidental release of nuclear radiation" were seen as widespread hazards. Again examples of localised hazards varied from one grid to another.

A "controllability" dimension was also found in five grids. In four of these grids, "earthquake" was seen as being outside one's control whereas "accident in the home", "rock-climbing accident" and "home fire" were seen as being within a person's control in three instances. A number of elements were construed as additional examples of either controllable or uncontrollable events in one or two grids.

The third dimension which was found in five grids bore the label "number affected". In three of these similarly labelled dimensions "accidental release of nuclear radiation" was construed as affecting a large number of people. As in other dimensions, a number of elements were found at either pole in one or two grids.

Even if contradictory ratings, caused either by differences in construing or by ratings of zero, are taken into account, it nevertheless remains that the comparisons of the elements contained in similarly labelled dimensions yielded weak results. As these results may indicate differences in construing as well as weaknesses in labelling, they subsequently served only as complementary information to the results of construct comparisons.

10.1.5 Comparisons of constructs

As pointed out earlier, all elements were common to all individual SA grids. However, the constructs were different from one grid to another. Therefore, comparisons based on constructs could not be as systematic as those based on elements. It was nevertheless possible to apply the same methods to the comparisons of constructs which were applied to the comparisons of elements. Firstly, a comparison was performed to find out whether constructs which appeared in more than one individual grid belonged to similarly labelled dimensions. Secondly, the contents of similarly labelled dimensions were examined to see whether their constructs revealed some constants. This section is the description of these two types of comparisons.

The first analysis necessitated an examination of the extent to which constructs were comparable across individual grids. At first sight, because of differences in wording, there appeared to be very few constants among the 358 constructs used in the 17 grids. A closer scrutiny revealed that there were only 20 unique constructs, similar to no other construct within their own grid or in other grids; these constructs were found in ten grids. The remaining 338 constructs could be reduced to 44 constructs found in two or more grids. These common constructs are listed in Table 10.2. This table also mentions the number of grids in which these constructs were found. Quite a few of these constructs were repeated within some grids. Therefore, the number of times that a construct was identified is also mentioned in Table 10.2; because of repetitions within grids, this number is greater than the number of grids in 34 cases. The categories under which common

TABLE 10.2

LISTING OF COMPARABLE CONSTRUCTS IN SA GRIDS

CONSTRUCT	NUMBER OF GRIDS	NUMBER OF CONSTRUCTS
<u>Origin of danger</u>		
Natural/man-made	12	23
Human cause/ no human cause	8	10
Blame assignable/ no blame assignable	3	3
Self responsible/ self not responsible	5	6
Internal/external	2	6
<u>Characteristics of hazards</u>		
Necessary/unnecessary activity	8	18
Occupational/ not occupational	6	9
Potential/ present	2	2
Near/ far	4	4
Moving/ stationary	3	3
Slow/ fast event	3	3
Specific/ non specific location	2	3
Open/ enclosed	3	5
Large/small concentration of people	3	3
<u>Threat</u>		
Frequent/ infrequent occurrence	2	6
High/ low risk of accident	8	12
Most dangerous/ least dangerous	4	8
Safe/unsafe	3	5
Sudden/ continuous threat	2	2

TABLE 10.2 (continued)

CONSTRUCT	NUMBER OF GRIDS	NUMBER OF CONSTRUCTS
Consequences		
Major/ minor	9	27
Large/small consequences	11	17
Fatal/ survivable	6	10
Many killed/ few killed	5	6
Many affected/ few affected	8	11
Personal/ impersonal	9	12
Instantaneous/ long-term consequences	4	4
Reversible/ irreversible	3	3
Painful/ painless	2	4
Human intervention		
Own control/ out of own control	11	24
Rely on others/ rely on self	3	4
Avoidable/ unavoidable	6	10
Preventable/ unpreventable	3	4
Precautions/ no precautions	2	2
Foreseeable/unforeseeable	2	5
Easy/ difficult to escape	7	12
Reactions		
Aware/ unaware of danger	3	5
Sleeping/ awake	3	4
Familiar/ unfamiliar	4	6
Ugly-hideous/ not ugly	2	6
Scaring/ not scaring	5	7
Worry-concern/ non-worry, unconcern	3	7
Acceptance/ non acceptance	4	5
Panic-chaos/ orderly-calm	3	7
Public reaction/ no public reaction	4	5
Miscellaneous (unique)	10	20
Total		358

constructs are grouped do not reflect any particular aspect of the grids or of sample tendencies. They were chosen arbitrarily and introduced, along with a listing not taking into account the order of importance, in order to facilitate the comprehension of sometimes cryptic construct labels.

The review and comparison of dimension to which they belonged was performed for all 44 common constructs. However, the results of this comparison are described below only for the 22 constructs found in four or more grids. This was done for two reasons. Firstly, a full description would have made extremely tedious reading. Secondly, for some constructs found only in two or three grids, some wording differences were overlooked in the reduction process. These overlooked differences could have biased the *comparisons of dimensions*. In contrast, the reduction process was a lot more rigid for the more frequently identified constructs, thus diminishing the risk of bias. Those constructs which had been excluded from reprocessed grids because of perfect correlations have been re-inserted for the purpose of the comparisons described below. Another detail is necessary for the understanding of these comparisons: some constructs loaded significantly on more than one dimension.

The most frequently elicited construct was "natural/man-made"; twelve respondents mentioned this construct a total of 23 times. Of the 38 significant loadings for these constructs, ten were in the six dimensions labelled "natural/man-made". Six of the loadings were found in the three "enclosed/open-air" dimensions; five were found in three of the five dimensions labelled "widespread/localised"; five more loadings were encompassed in two of the five "controllability" dimensions. The remaining twelve loadings were scattered among seven other dimensions.

Two constructs were elicited by eleven respondents. They were "large/small consequences", mentioned 17 times, and "own control/out of own control", mentioned 24 times. Of the 20 significant loadings for "large/small consequences", eight were found in three "natural/man-made" dimensions and in three "controllability" dimensions; the remaining six loadings were encompassed in four dimensions all related to the concept of severity of consequences, such as "number affected". Five of the 30 significant loadings of "own control/out of own control" appeared in four out of the five "controllability" dimensions. Two "likely/unlikely" dimensions encompassed an additional eight loadings. The majority of the remaining 21 loadings were related to eight dimensions referring either to the concept of the scope of human intervention (natural/man-made, preventability, etc.) or to the concept of severity of consequences (severity, widespread/localised, etc.).

Two other types of constructs were items of nine grids. Firstly, "major/minor" was mentioned 27 times and accounted for 29 significant loadings. Nearly half (14) of these loadings were in four of the seven "severity of consequences" dimensions; the remaining 15 loadings belonged mostly to dimensions referring either to potential consequences in one form or another (e.g. number affected, instantaneous/long-term, etc.) or to the repulsiveness of hazards (ugliness, hideousness, horror). Secondly, there were twelve mentions of "personal/impersonal"; this construct was a significant member of 17 dimensions, no one dimension encompassing more than four of the loadings.

However, this construct tended to be associated with the concept of population at risk and with the assessment of the extent of potential devastation (severity of consequences, number affected, widespread/ localised).

The constructs "human cause/no human cause", "necessary/ unnecessary activity", "high/low risk of accidents" and "many/few affected" were each mentioned by eight respondents. The first of these constructs, "human cause/no human cause", displayed no distinct affinity since no dimension encompassed more than three of its significant loadings. The majority of these 18 loadings were found in two types of dimensions: those related to the scope of intervention (voluntariness of activity, natural/man-made, etc.) and those reflecting the severity of potential consequences (number affected, severity). The second construct, "necessary/unnecessary activity", was mentioned 18 times for 22 significant loadings. Its most important affiliation was to the "voluntariness of activity" dimension; eight of the loadings were found in six of the eight dimensions thus labelled. The remaining 14 loadings related this construct to eight dimensions again reflecting the scope of human intervention and the severity of consequences. It was hardly surprising that six of the 16 significant loadings for the "high/low risk of accident" were found in the three dimensions labelled as "likely/unlikely". The other ten loadings were found in seven other dimensions, most of which were akin to the concept of severity of consequences. As for the fourth construct, "many/few affected", of the 19 significant loadings four were encompassed in each of three dimensions: "number affected" in four grids (out of the five from which such a dimension was extracted), "widespread/localised" in three grids and "controllability" in two grids. Six dimensions, most of which were akin to the notions of severity of consequences and of scope of human intervention, attracted the remaining seven loadings.

"Easy/difficult to escape" was the only construct mentionned by seven respondents. The 12 constructs loaded significantly 16 times, but never more than three times in any given dimension. Of course, the only dimension labelled "easy/difficult escape" included one of these constructs. The remaining 15 loadings were scattered among nine dimensions around the usual two themes.

Three constructs were mentioned by six respondents each. They were: "occupational/not occupational", "fatal/survivable" and "avoidable/unavoidable". The "occupational/not occupational" construct was found twice in the only dimension labelled "occupational/not occupational". It is interesting to note that four "voluntariness of activity" dimensions each included one significant loading of this construct. The remaining loadings were found in much the same type of dimensions as in the case of the preceding constructs. The second construct, "fatal/survivable", was thirteen times a significant member of a dimension. Four of these loadings were found in one dimension labelled "horror", whereas three dimensions labelled "severity of consequences" included this construct among its members. The remaining six loadings did not show any particular tendency in five diverse dimensions. The third construct which appeared in six grids, "avoidable/unavoidable", elicited fourteen significant loadings among its ten mentions. Two of these loadings belonged to two "natural/man-made" dimensions and three to one "controllability" dimension. Nine differently labelled dimensions each generated one of the remaining significant loadings.

As the subsequent constructs were mentionned less and less frequently clustered affinities with specific dimensions became more difficult to identify. Thus, the three constructs mentionned only by five respondents, i.e. "self responsible/self not responsible", "many/few killed" and "scaring/not scaring", could not display any large concentration of loadings in any dimension. "Self responsible/self not responsible" loaded three times among two "number affected" dimensions, and once each in "voluntariness of activity" and "nature of consequences" dimensions. As for "scaring/not scaring", its five loadings were distributed in the following way: two in a "severity of consequences" dimension, two in a dimension labelled "horror" and one in a dimension under the heading "widespread/localised"; two of these constructs were not assigned to any dimension.

The six constructs found in four grids are of some interest only because three of them had labels also used to describe dimensions, albeit infrequently extracted dimensions. The first of the six constructs, "near/far", loaded significantly once in each of the fol-following dimensions: "modern/traditional", "widespread/localised", "enclosed/open-air" and "severity of consequences"; one more construct was found in an unlabelled dimension. The second construct was "most dangerous/least dangerous" and accounted for eight significant loadings; this construct is possibly the only one among those found in four grids whose loadings displayed some constancy. Five loadings were found in the dimension labelled "hideousness", one was found in a "riskiness" dimension, one also in the "horror" dimension . As for the "instantaneous/long-term effects" construct, one of its eleven loadings was found in the only "short-term/long-term effects" dimension. The other significant loadings were distributed as follows: two in as many "likely/unlikely" dimensions, and one each in "controllability", "foreseeability", "number affected" and "voluntariness of activity" dimensions. The fourth construct, "familiar/unfamiliar", loaded once in each of the two dimensions labelled similarly. Two "severity of consequences" dimensions also accounted for one significant loading each; the only other significant loading was in a "widespread/localized" dimension, other affiliations of this construct not being definable. The "acceptance/non-acceptance" construct loaded twice in a "natural/man-made" dimension, and once in one of each of the following dimensions: "acceptance of risk", "riskiness", "horror" and "controllability"; furthermore, one significant loading was in an unlabelled dimension. Finally, the "public reaction/no public reaction" construct was not assigned to any dimension in one grid; it was a significant member of two "severity of consequences" dimensions as well as of one "modern/traditional danger" and one "number affected" dimension.

One of the important observations emerging from this first series of comparisons was that there was a close concordance between similarly designated constructs and dimensions. To a certain extent, this is not very surprising, since the labels affixed to dimensions were inspired from the constructs they encompassed. But the concordance between constructs and dimensions was such that it suggested that on the whole, the labels were fairly apposite. The second series of comparisons was necessary, however, before the appropriateness of labels could be confirmed. Other implications of these comparisons are analyzed in section 10.3 of this chapter.

The second series of comparisons consisted of an analysis of the constructs comprising the contents of the most frequently identified dimensions. As pointed out in the discussion on the comparison of elements, "voluntariness of activity" was used to describe eight dimensions. Of these eight dimensions, six included at least one "necessary/unnecessary activity" construct. Four dimensions (including the two which did not encompass a "necessary/unnecessary activity" construct) each included an "occupational/non occupational" construct. The eight dimensions also attracted a total of 19 other constructs under 12 headings.

The "severity of consequences" dimensions were among the largest throughout the sample. Apart from the dimensions reflecting the repulsiveness of hazards ("horror", "hideousness"), they were comprised of an average of more than ten constructs per dimension. Along with the mean rank of extraction of these dimensions, this abundance of constructs is another indication of the prominence of this concept in respondents' mind. Six of these seven dimensions attracted a total of eight "large/small consequences" constructs. Four dimensions encompassed 14 "major/minor" constructs between them. Three dimensions were associated with the construct "fatal/survivable". The construct "many/few killed" was found a total of three times in two dimensions. A total of 29 construct headings explained the other 43 significant loadings, including at least one mention of all the other constructs related to the various aspects of potential consequences (e.g. "painful/painless", "many/few affected", etc.). It is interesting to note that seven of the 20 unique constructs (i.e. found only once in one grid, such as "black/white", "luck/destiny", etc.) were found to load significantly on some of these dimensions.

All six of the "natural/man-made" dimensions included at least one similarly labelled construct; two dimensions attracted two of these constructs and one dimension encompassed three of them. Other noteworthy constructs included "human cause/no human cause" (three constructs in two dimensions), "large/small consequences", (one in each of three dimensions), and "own control/out of own control" (four in three grids). The other 21 significant loadings in these dimensions were accounted for by 18 constructs, including five of the 20 unreplicabled constructs.

Five dimensions received the label "widespread/localised". Three of these dimensions attracted a total of five "personal/impersonal" constructs. Three dimensions included a total of four significant loadings by "many/few affected" constructs. Four "enclosed/open" constructs were also found in two dimensions. Furthermore, it is interesting to note that in three dimensions natural hazards were construed five times as being widespread and man-made hazards as being localised. The remaining 26 significant loadings were distributed among 19 constructs.

Of the five "controllability" dimensions, four included a total of six "own control/out of own control" constructs. All the other facets of the concept of scope of human intervention were also represented in these dimensions: "preventable/unpreventable" three times in two grids, "natural/man-made" five times in three grids, "foreseeable/unforeseeable" three times in one grid, "human cause/ no human cause" and "self responsible/self not responsible" once in one grid. Eleven other constructs accounted for the remaining 17 significant loadings.

The last dimensions discussed in this analysis were labelled "number affected" and were extracted from five grids. In four of these dimensions were found a total of five "many/few affected" constructs. One "personal/impersonal" construct was found in three dimensions. Eighteen constructs accounted for the remaining 29 significant loadings.

Pursuing these comparisons to those dimensions found only in three grids or less would bring to light very little additional information. Two important observations emerged from these comparisons based on constructs. Firstly, it was noticed that a majority of dimensions incorporated a number of constructs representing secondary characteristics of their core. Quite a few of these secondary constructs were shared by differently labelled dimensions. At the end of this chapter it is argued that an overlap in terms of shared constructs and elements indicates that some dimensions are closely inter-related. Shared constructs may also be an indication that some dimensions are sub-sections or specific facets of other, more general dimensions. Such fragmentation of dimensions is discussed in Chapter 11.

Secondly, the second series of comparisons seemed to confirm that the contents of similarly labelled dimensions were, in their essence, very similar. The comparisons based on constructs removed most of the doubt raised by the comparisons based on elements. The analysis of constructs showed that constructs tended to belong to their homonymous dimensions, and that dimensions were generally elaborated around similarly designated constructs.

During these analyses of constructs it was noticed that some construct loadings, like element loadings discussed earlier, were contradictory. Since on the whole constructs were used in the same way across grids, the main source of the observed inversions turned out to be the use of zero as rating for certain elements. The impact of the use of zeros is discussed in this next section.

10.1.6 Zeros or threes

In the SA sample, respondents were given the possibility to signify that some constructs were irrelevant in the assessment of certain elements, in other words, that there were elements outside the range of convenience of some constructs. Respondents who wanted to indicate that a construct was irrelevant for an element were instructed to give a rating of zero. Table 7.10, in Chapter 7, indicates that six respondents did not avail themselves of the opportunity of assigning zero ratings. In the other eleven grids, the number of zeros varied between 9 and 304.

On the basis of a high number of zero ratings in some grids, many inversions and inconsistencies were expected. Indeed, the inversions of element and construct loadings which occurred were mostly caused by large numbers of zeros. It was noticed, however, that not all constructs having elicited many ratings of zero behaved inconsistently. Whether or not many zeros led to inversions depended on the nature of constructs.

Two types of constructs were found in SA grids. The first type of constructs consisted of a characteristic whose presence in an element was rated as five and whose absence was rated as one. For instance, on the "many/few killed" construct elements which were assessed as potentially killing many people were rated as five. This type of construct did not lead to any inversion. The second type of constructs was truly bi-polar since the ends of the constructs were symmetrical opposites (e.g. "white/black"). The inversions were all found in this type of constructs.

It would seem that there were inversions only in the second type of constructs because ratings of zero may have been more "natural" in the first type of constructs. For instance, if one element was judged by a respondent not as being a hazard, assessing the number of people who could be killed became irrelevant. Thus, a rating of zero, used in order to signify that the construct is irrelevant, could at the same time be interpreted as the natural continuation of the rating scale. In other words, a rating of zero could be interpreted as a statement that no one could possibly be killed, as a rating of 1 is considered as meaning few deaths. In contrast, on the bi-polar constructs, the neutral point should in fact be in the middle of the scale; on a scale from one to five, the neutral rating should be three. Because zero is the mathematical prolongation of a scale from five to one, using zero as the neutral point is equivalent to artificially placing more emphasis than usual on one end of the bi-polar constructs.

In order to test the validity of this explanation, some grids containing zeros were reprocessed after the ratings of zero had been changed for threes. Following such reprocessing the previously inverted constructs straightened up, whereas some constructs of the first type described above behaved in a strange way. Thus, it would appear that the choice of a neutral point should take the type of construct into consideration.

These observations about the choice of a neutral point raise the question of the extent to which the use of zeros introduced bias in the results. The very good robustness of the dimensions discussed so far in this chapter is a first indication that the amount of bias created by the use of zeros was relatively small. The second similar indication comes from the fact that most constructs which elicited ratings of zero were not truly bi-polar; for these constructs it would have been the use of three as the neutral rating which would have introduced bias, as illustrated by the reprocessing of grids described above. Thirdly, in order to confirm these hypotheses some grids were reprocessed after their constructs containing zeros were removed. The nature and the order of importance of their main dimensions were practically unchanged. The contents were also unchanged except for the omission of the removed constructs.

So far, mostly technical aspects of the SA results have been presented and discussed. Needless to say, the interpretation of these results is very important. However, it was found that contrasting and comparing these dimensions with those identified in the fourth sample raised important questions to be discussed in the next chapter. Therefore, the discussion of results is done in section 10.3 of this chapter, after the description of PH results in this next section.

10.2 RESULTS FROM THE FOURTH SAMPLE

The fourth sample was rather limited. The only three respondents were experienced speleologists who generated a list of nine pot-holing hazards. This list of elements was presented in Table 6.3, and for reading convenience it is also in fold-out form in Appendix 6. From these elements, six constructs were elicited in a group discussion. These constructs were listed in Table 6.3 and can be found in Appendix 7.

10.2.1 Individual grids

The three PH grids were processed as usual by Ingrid 72. Then cluster analysis and principal component analysis were performed.

Then, $\bar{X}$ and $\bar{Z}$ grids were computed and processed. These are discussed in the next section of this chapter. This section is mainly concerned with individual grids. However, as this fourth sample is quite small, some aspects of the $\bar{X}$ and $\bar{Z}$ grids are discussed along with the corresponding aspects of the three individual grids.

10.2.1.1 Number of components

One of the first points which deserves consideration is the number of components extracted by principal component analysis from the individual grids and from the $\bar{X}$ and $\bar{Z}$ grids. In all five grids, the Bartlett test was found to give negative results.

It must be pointed out that the Bartlett test is initiated by including two major components; another component is included with each following step of the test. According to Slater (1972), if there are only one or two significant components in a grid, the test is bound to give negative results since the significant components are overlooked from the start.

Thus, it would appear that in all grids, there was only either one or two significant components. As there were only six constructs in each grid, there could only be six components at most. Therefore, it is not surprising that there would be fewer significant components in PH grids than in grids from other samples.

It was necessary, however, to ascertain how many dimensions could justifiably be interpreted. Indications of the number of significant dimensions were sought from the number of components on which at least one construct or element loaded significantly, and from the number of clusters in each grid. Cattell's Scree test and Kaiser's criterion of a latent root of 1 were also applied. The details of these four indicators are listed in Table 10.3.

In this table, one can notice that Kaiser's criterion always yielded a smaller number of components than did the Scree Test. This is explained by Child (1976) when he says that for small numbers of items (i.e. constructs) Kaiser's criterion may prove more reliable than the Scree test. And indeed, the results of a latent root of 1 are more in line than the results of the Scree test with the results from the other two indicators.

In two grids, only one cluster was extracted; in three grids constructs and elements were found to load significantly on only one component. However, there was always good agreement between the first cluster and the first component in each grid. Never were two clusters encompassed in one component or two components in one cluster.

It must also be remembered that, as there were only nine elements, cut-off points for both cluster analysis and principal component analysis were relatively high. Thus, the minimum correlation for a cluster of only two constructs, which was also the minimum loading for the first component, was ± 0.666.

TABLE 10.3

INDICES OF THE NUMBER OF DIMENSIONS

GRID	NUMBER OF CLUSTERS	NUMBER OF DIMENSIONS WITH SIGNIFICANT LOADINGS	SCREE TEST	LATENT ROOT OF 1
PH01	2	2	4	2
PH02	1	2	5	2
PH03	2	1	4	2
$\bar{X}$	2	1	3	2
$\bar{Z}$	1	1	3	2

So, it could be that because of high cut-off points, trends found significant by one analysis were not retained by the other analysis. Therefore, it was decided that for all but the $\bar{Z}$ grid, two dimensions would be retained for further discussion.

10.2.1.2 Constructs and elements

In order to understand the nature of the various dimensions, their contents had to be identified in terms of constructs and elements. As descrived earlier, cluster analysis and principal component analysis were performed on constructs in each grid.

A summary of the results of these two analyses appears in Table 10.4. The frequencies of allocation of each construct to each dimension by both analyses are also listed. Finally, the dimensions to which each construct was allocated by both analyses in both mean grids constitute the last part of the table.

Probably, the most noticeable feature of this table is that a first dimension seemed to account for a large portion of each individual and mean grid. In fact, such a first dimension explained from 55 to 70 percent of grid variances.

It is interesting to note that two components were found in two of the three grids whose latent root of the first component was lowest. Traces of a second significant dimension were not found only in that grid (the $\bar{Z}$ grid) which had the highest first latent root.

TABLE 10.4

SUMMARY OF THE EXTRACTION OF DIMENSIONS FROM PH CONSTRUCTS

CONSTRUCT	GRID						DIMENSION FREQUENCIES						MEAN GRIDS			
	PH01		PH02		PH03		1		2		NONE		$\overline{X}$		$\overline{Z}$	
	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F	Q	F
1	1	1	1	1	1	1	3	3					1	1	1	1
2	1	1	1	1	1	1	3	3					1	1	1	1
3	2	1	1	1	1	1	2	3	1				1	1	1	1
4	2	2		2	1	1	1	1	1	2	1		2	1	1	1
5	1	1	1	1	2	1	2	3	1				2	1	1	1
6	1	1		2	2	1	1	2	1	1	1		2			

Before interpreting these dimensions, their contents were defined further by identifying the elements which loaded significantly. The results of this analysis are listed in Table 10.5.

A remarkable feature of this table is that, in elements, there is no trace of a second dimension. In other words, no element could be identified which was related to the second dimension extracted from constructs in some grids. This may be due to a high cut-off point ; the minimum loading on the first component was 0.811, the highest cut-off point of all in this research.

10.2.1.3 Comparability of dimensions

Having defined the contents of each dimension in each grid, these dimensions were then labelled. It was found that the first dimension in each of the individual grids always contained the first ("controllable-uncontrollable") and the second ("lack of thought-bad luck") constructs and element 8 ("loose rocks in boulder choke"). In all three individual grids, the label "scope of intervention" could be affixed to the first dimension. In this respect the first dimension was very similar to that found in other samples.

As could be suspected from Tables 10.4 and 10.5, the individual grids were not as unanimous about the second dimension as they were about the first dimension. There was not one construct which could be found to belong to a second dimension in all grids; furthermore, in no grid could any element be found to load significantly on this second dimension.

TABLE 10.5

SUMMARY OF THE EXTRACTION OF DIMENSIONS FROM PH ELEMENTS

ELEMENT	GRID			DIMENSION FREQUENCIES			MEAN GRIDS	
	PH01	PH02	PH03	1	2	NONE	$\overline{X}$	$\overline{Z}$
1	1		1	2		1	1	1
2						3		
3						3		
4						3		
5		1	1	2		1	1	1
6						3		
7						3		
8	1	1	1	3			1	1
9						3		

The label "severity of consequences" can probably be affixed to the second dimension of grids PH02 and PH03; construct 6 ("likely to kill/unlikely to kill") is found in this dimension in both grids. As for the second dimension in grid PH01, one can tentatively use the label "nature of a danger" since construct 4 ("technical factor/physical factor") is the only one retained both by cluster analysis and by principal component analysis. However, the contents of this dimension are insufficiently precise for the label to be beyond doubt. The label "severity of consequences" does not seem apposite in grid PH01, since construct 6 belonged to the first dimension.

Thus, whereas the first dimension of all three grids appeared to convey the notion of "scope of intervention", in only two grids did a second dimension appear comparable.

10.2.2 Consensus dimensions

Unanimity about a first dimension and lack of coherence about a second dimension was reflected in the results of the three techniques of extraction of consensus dimensions. For instance, a first dimension could be extracted from both mean grids ($\bar{X}$ and $\bar{Z}$); however only in the $\bar{X}$ grid could a second dimension be identified.

The first consensus dimension closely resembles the first dimension in individual grids. Frequencies of assignation in cluster analysis and in principal component analysis, and clusters and components in both the $\bar{X}$ and the $\bar{Z}$ grids all agree that constructs 1 and 2 are the core of this dimension. Constructs 3 and 5 are also important members of this dimension although both show affinities with a second dimension.

The label "scope of intervention", chosen for the first dimension of all individual grids, appears apposite for the first consensus dimension. Construct 1 ("controllable/uncontrollable") has been encountered in other major dimensions in other samples. Construct 2 ("lack of thought/bad luck") conveys the idea of the extent to which the occurrence of a hazard is attributable to a person's intervention or action. And so do constructs 2 ("due to my party/due to another party") and 5 ("poor planning/poor moving"), to a certain extent. Therefore, the same important connotations are present in this dimension that were present in the corresponding dimension in other samples.

At first sight the nature of the second consensus dimension is rather ambiguous. Although traces of a second dimension were found in all three individual grids, these indications were sometimes less than robust. Furthermore, only in the $\bar{X}$ grid was there a second cluster; there was no second cluster in the $\bar{Z}$ grid and no second component in either mean grid.

The absence of a second cluster in the $\bar{Z}$ grid is probably due to five of the six constructs belonging to the first cluster. Construct 6, being the only one left out, had no other construct with which to correlate and hence to form the core of a cluster. But then, it may be important that it was construct 6 ("likely to kill/unlikely to kill") which was left out.

A second look at the loadings of the second dimension in both mean grids revealed an interesting detail. In both grids construct 6 accounted for the highest loading (.61 and .62) on the second dimension. Such loadings were not considered statistically significant in this sample, but they would have been considered significant in any other sample.

Only in grid PHO2 was construct 6 not assigned to any cluster. That was because its correlation with construct 4 (the only other construct left out of the first cluster) was -0.553. There again this correlation was not significant enough to justify a second cluster; however in any other sample the same correlation would have been found significant.

These loadings and correlations are indirect indications that a second dimension probably revolves around construct 6. More direct evidence is found in the second cluster of the $\bar{X}$ grid; this cluster encompasses construct 6 as one of its significant members.

Thus, "severity of consequences" could well be the idea conveyed by the second consensus dimension. This would be in line with findings from the first two samples. However, the evidence from this fourth sample is tenuous; a second dimension was not always found to be significant.

In the second (UG) sample, there were indications that the concepts of "scope of intervention" and of "severity of consequences" were correlated. Similar indications were found in this fourth sample.

For instance, construct 6 was found in the first cluster in grid PH01 and in the first component in grids PH01 and PH03. In addition, construct 4 and 5 sometimes belonged to the same dimension as did construct 6, and sometimes to a different dimension (e.g. in the components of the mean grids).

In all cases, these indications are consistent with those found in the second sample. That which is perceived to be beyond one's control is also perceived as having severe potential consequences. Conversely, that which is perceived to be within one's control is also perceived as having mild potential consequences.

Very little additional information can be extracted from the results of this fourth sample at this stage. However, some interesting points are raised in the next section, where the results of the third and fourth samples are discussed.

10.3 DISCUSSION OF THE RESULTS

10.3.1 Interpretation of SA results

There are a number of observations and results from SA grids which deserve a very close scrutiny. In this respect, the first point, and possibly the most crucial one, which arouses curiosity is the apparent paradox of relatively robust dimensions not unanimously extracted from all individual SA grids. It may be necessary to cast a second look at the robustness of the dimensions in the first place, and to attempt to understand why dimensions were never common to more than eight grids out of 17.

The affiliations of constructs to dimensions were not all reviewed in this chapter. Nor were the constructs contained in all comparable dimensions analysed. Although the results were not presented in this chapter, the less frequent constructs and dimensions were also subjected to careful scrutiny and revealed very much the same consistency and coherence as did the more frequent constructs and dimensions.

It can be argued that there are indications which jeopardize the coherence of some comparable dimensions. For instance, two of the eight "voluntariness of activity" dimensions did not include a "necessary /unnecessary activity" construct. However, it must be pointed out that these two dimensions both included an "occupational/not occupational" construct whose "not occupational" pole in fact read as "leisure". In addition, these two dimensions belonged to grids which did not contain a "necessary/unnecessary activity" construct. Furthermore, most prominent constructs in these two dimensions found their counterparts in one or more of the other six similarly labelled dimensions. The same type of arguments can be used for practically all the other dimensions which did not encompass their homonymous construct. Therefore, as fas as constructs are concerned, it seems difficult to question the robustness of comparable dimensions.

One apparent weakness of the dimensions appears to be that the analysis of the behaviour of elements did not reveal any major trend. It must be pointed out that to a certain extent, there were small clusters of elements which showed some affinity with some common dimensions. But the weakness of these clusters of elements does not necessarily cast a serious doubt on the robustness of common dimensions.

As pointed out earlier, some elements were assessed differently by different respondents. Thus, the labelling of dimensions from their elements would be more subjective than it is when done from their constructs, since one would have to impose one's own construing system upon the results in order to understand why certain elements behave the way they do. Thus, the content validity must be judged from the constructs rather than from the elements. It was argued earlier that the analysis of constructs revealed very good comparability of common dimensions.

It is somewhat surprising, then, that contrary to what happened in other samples, no dimension appeared to be common to more than half of the grids. In fact, the main clue to the answer seems to be in the difference between the labels used in the SA sample and those used in the other three samples. Comparisons of these labels reveal that in the other three samples, the main labels (e.g. "scope of human intervention", "severity of consequences") are fairly general; in contrast, the labels used in the SA sample are more specific (e.g. "preventability", "number affected"). It is as though the broader dimensions of other samples were fragmented into more specific facets in the SA sample.

It is argued in the next chapter that this fragmentation is due to the very nature of elements used in the SA sample. But for the time being, it is interesting to examine this fragmentation a little more carefully. Without it being explicit, in other samples the concept of "scope of human intervention" encompassed two important concepts: the extent to which the presence (or the origin) of a danger was attributable to some human intervention, and the notion of controllability of a hazard. In turn, "controllability" was a broad heading for at least three concepts, the first being the extent to which human intervention can prevent a hazard from becoming an accident, the second being the possibility that some human intervention can reduce or eliminate the consequences of a hazard leading to an accident, and the third being the possibility for a person to control his/her exposure to a danger. All these various facets of the concept of scope of human intervention are present in SA dimensions. For instance, human intervention in the origin of danger is illustrated in the six "natural/man-made" dimensions. Five dimensions were also labelled "controllability"; but, because the "own control/out of own control" construct was often associated with "likely/unlikely" and with "avoidable/unavoidable" in these dimensions, the main facet of controllability depicted by them appears to be the control over a hazard leading or not to an accident. The notion of control over one's own exposure to a hazard is the main theme of the eight "voluntariness of activity" dimensions, where "necessary/unnecessary activity" is strongly correlated with "own control/out of own control". Finally, the notion of control over consequences is reflected by the three "preventability" dimensions which encompass "own control/out of own control" and various constructs related to the extent of potential consequences as their most prominent members.

The concept of severity was also fragmented into more specific dimensions. There were seven dimensions labelled "severity of consequences" among the seventeen grids. But there were also five dimensions labelled "number affected", and five "widespread/localized" dimensions; the labels "size of population at risk", "short-term/ long-term effects", "nature of consequences" and "extent of material damage" were also used once each. The concepts of the type of consequences (material destruction only or lives lost), the extent of material damage, the dispersion of destruction, the number of people at risk or affected, the amount of time between the occurrence of an accident and the appearance of its consequences all seemed related to the concept of severity of consequences. The dimensions labelled "severity of consequences" rested mostly on constructs such as "major/ minor" and "large/small consequences". The more specific dimensions were based on more specific constructs, but most of them also included one or more "major/minor" or "large/small consequences" constructs.

If all the specific dimensions are reallocated to the general dimensions they were derived from, then at least one "scope of human intervention" would be found in thirteen grids and at least one "severity of consequences" would appear in eleven grids. But the mere fact that two or more general dimensions with the same label would be found in some grids is sufficient justification to prohibit reverting to the general labels.

In view of what was found in the first two samples this fragmentation is of the utmost importance. Therefore, it is discussed in further detail in the next chapter on comparisons between samples.

Another noteworthy feature of this sample is that in three grids there was only one significant dimension explaining an average of four fifths of the common variance. Two of these unique dimensions center around the repulsiveness of hazards ("hideousness" and "horror") and the third is labelled "severity of consequences". In addition to being unique in their respective grids, these three dimensions had other characteristics in common. They all emcompassed practically the totality of constructs in their own grid for an average of nearly 18 constructs each. Upon examination, it can be observed that the majority of constructs in all three dimensions are of two types: some referring to the repulsiveness of hazards and some concerning one aspect or another of the severity of potential consequences. These are the only three dimensions in the whole sample whose construct contents are thus arranged. It is interesting to note that the only grids where hazard repulsiveness is the subject of some constructs are also the only grids which yield only one dimension. This would seem to suggest that for three respondents an emotional reaction inspired by hazard repulsiveness obliterated a rational assessment of hazards; in the other fourteen grids, a relatively articulate assessment of hazards seems to have been carried out. It must be pointed out that in grid SA02, from which an "ugliness" dimension was extracted, no construct referring to the repulsiveness of hazards was found; the label of this dimension is based only on one construct and two elements, and there is a possibility that this label may not be the most appropriate one for this dimension.

The fact that most grids yielded fragmented dimensions does not mean that these specific dimensions are totally independent from each other. In fact, there are indications that some of these dimensions are inter-correlated. For instance, it was pointed out earlier that many dimensions depicting facets of the scope of human intervention had some constructs in common (e.g. "own control/out of own control"). The same can be said of dimensions referring in one way or another to the concept of severity of consequences; these dimensions share constructs such as "major/minor" and "large/small consequences". On the basis of such common constructs, a network of inter-correlations can be identified. In general, "natural" hazards are deemed "not controllable" and thus potentially leading to very "severe consequences"; in contrast, "man-made" hazards are assessed as "controllable", and their potential consequences tend to be viewed as trivial. These associations are confirmed by high correlations in grids where constructs reflecting these concepts are found. Furthermore, "natural" hazards are generally assesed as "widespread" and as repulsive, whereas it is believed that "man-made" hazards are "localised" and merely "unpleasant". These observations are in line with indications in other samples that "controllability" and "severity of consequences" are correlated. This is discussed further in Chapter 11.

10.3.2 Interpretation of PH results

In view of the discussion above, it seems that the label "scope of human intervention" was apposite for the main dimension extracted from all PH grids. This dimension appears to be a general rather than a specific dimension. Its most frequent constructs ("controllable/uncontrollable", "lack of thought/bad luck", "due to my party/due to another party", "technical factor/physical factor") refer to various facets of the origin of a danger as well as of the reasons why this danger should lead to an accident.

The fluctuating nature of a second dimension, the difficulty in establishing whether it is significant and the small number of constructs and elements comprising it when it is significant make it difficult to decide whether it is a general or a specific dimension.

The label "severity of consequences" is apposite only if the dimension is a general one. The main construct of this dimension, however, appears to refer to a specific aspect of the concept of severity of consequences, namely likelihood of fatality.

The fact that a second dimension was difficult to identify in some instances appears to be related to the size of the grid used by PH respondents. Nine elements rated only on six constructs probably generated little common variance; a small amount of common variance is usually difficult to apportion, especially because cut-off points have to be high. In those instances where only one dimension was identified (e.g. in $\bar{Z}$ grid), the possibility of an emotional dimension such as that which was found in three SA grids was contemplated. However, because of the "scope of human intervention" aspect common to five of the six constructs, this possibility of an emotional factor was discarded as being unlikely.

The changing nature of a second dimension across PH grids suggested that there was not a comparable second dimension in all grids. Thus, it would seem that PH grids yielded specific rather than general second dimensions. Because of very limited contents, it was difficult to define precisely the subtle differences between these second dimensions. Nevertheless, those constructs which belonged sometimes to a first dimension and sometimes to a second dimension were further indications that in general, the notions of "scope of human intervention" and "severity of consequences" were not totally independent.

10.3.3 Summary

Two very different samples were discussed in this chapter. Because SA respondents used grids whose elements were identical but whose constructs were different, and because PH respondents rated a very small grid, the analysis of these two samples could not be done as systematically as the first two smaples. However, some very important observations arose from these analyses.

Firstly, it was noticed that in some instances, general dimensions can be broken into more specific concepts. However, it seems that such an articulate method of construing can be overridden by an emotional reaction to the repulsiveness of hazards. But even when specific rather than general concepts are used, it is possible to notice the affiliation of the specific concepts to the more general ones. Furthermore, some affinities between the broad concepts could be observed. All these observations have an important bearing on the comparison of results between samples.

CHAPTER 11 : COMPARISONS OF RESULTS BETWEEN SAMPLES

The four samples reviewed in the previous chapters differed in a number of ways. Many personal characteristics of the respondents were different between samples. The grids used by respondents differed as did the ways in which elements were chosen, constructs were elicited and grids assembled and scored. These differences were arranged so as to make it possible to test for the possible influence of the personal characteristics and of the various methodological variables on the identified hazard assessment structure. The outcome of these tests is discussed in this and in the last chapters.

When the results of the processed grids were reviewed, some similarities as well as a number of discrepancies were noticed. The first part of this chapter is a description of comparable results between samples. This description rests mainly on individual rather than consensus grids for two reasons. Firstly, it was argued in the previous chapters that some consensus dimensions (numbers 3 and 4 in the GR sample, 4 and 5 in the UG sample and 2 in the PH sample) may have been inadequate representations of individual dimensions. Secondly, no consensus dimensions could be calculated in the SA sample. Therefore, individual results provided a better basis for comparisons.

The second part of this chapter is a discussion of differences in results, including some tentative explanations for the existence of these dissimilarities. Differences in respondent characteristics and in grid techniques are examined in an attempt to determine whether these variables explain the observed differences. Finally, the most important dimensions are reviewed across samples. In order to do so, consensus as well as individual dimensions are discussed.

11.1 SIMILARITIES BETWEEN SAMPLE RESULTS

When the hypotheses of this research were formulated, it was predicted that the main dimensions in all samples would be similar. It turned out that similarities between samples were fewer than expected. Only two methodological features of the results emerge as similar when the four samples are compared. Firstly, within all samples, there tended to be close agreement between the results of cluster analysis and those of principal component analysis. In most of the cases where discrepancies arose, it was either because a component had been fragmented into two clusters or because two components had been extracted as clusters in a reverse order.

Secondly, in all samples the Bartlett test tended to reveal a number of significant components which was greater than either the number of clusters or the number of components which could be labelled. These inadequacies of the Bartlett test have been discussed in Chapters 8 and 9.

To a certain extent, the nature of the main consensus dimension (almost unanimously matched by the corresponding individual dimensions) in three samples (GR, UG and PH) was labelled "scope of human intervention". The fact that it was identified as prominent by both data reduction methods indicates that the concept accounted for the largest portion of common variance in the grids. In turn, this suggests that "scope of human intervention"is the respondents' main concern when they assess hazards. It was mentioned in Chapter 5 that one of the assumptions underlying the grid method is that the structure identified within a grid serves as a basis for a respondent's behaviour. This research was not designed to test this assumption. Nevertheless, future research which would look into the influence of the concept of "scope of human intervention" upon behaviour in the face of danger might provide valuable information to those concerned with safety training and propaganda.

The second important consensus dimension was also a relatively accurate representation of the corresponding individual dimensions in the same three samples (GR, UG and PH). Under different labels (i.e. "dreadfulness", "dangerousness" and "severity of consequences"), this second dimension referred to the concept of severity of potential consequences. With hind-sight the "dreadfulness" label affixed to the second dimension in the GR sample does not appear apposite. It was pointed out in the previous chapter that the notion of dread was associated with the notions of horror and repulsiveness of hazards in three SA grids which yielded only one dimension each. The contents of these SA dimensions do not correspond to those of the second GR dimension. The latter does not contain any construct related to the notion of dread. As pointed out later, however, the contents of the GR dimension correspond to those of the second UG and PH dimensions. Thus, the label "severity of consequences" appears to be a more faithful description of the nature of the second GR dimension.

Various aspects of "scope of intervention" and of "severity of consequences" are also present in SA grids. There were 22 dimensions referring to one of the facets of "scope of intervention" and 21 dimensions related to some aspect of "severity of consequences".

Both concepts were represented by a total of 94 constructs throughout the 17 SA grids. Therefore, at first sight, there are few indications of which concept was prominent over the other.

The order of extraction of the various dimensions provides further information on this point. In six grids a dimension related to "scope of intervention" was extracted before a dimension referring to the concept of "severity of consequences". In contrast, in eight grids, a dimension related to "severity of consequences" was extracted before a dimension referring to the other concept; in one grid, "severity of consequences" was the label of the only dimension extracted; as pointed out in Chapter 10, the two unique dimensions reflecting the repulsiveness of hazards had important undertones of "severity of consequences".

Thus, there are no systematic indications of the prominence of either concept over the other, although the concept of "severity of consequences" takes precedence in a few more grids. This may indicate that, when major hazards such as SA elements are assessed, "scope of human intervention" and "severity of consequences" are equally important concerns in respondents' minds. Alternatively, this may suggest that for some respondents "severity of consequences" is the main concern. It was mentioned in Chapter 9 that this concept also took precedence in two of the eleven UG grids. This topic is discussed again later in this chapter.

It was pointed out in Chapter 10 that SA dimensions focused on specific aspects of the two concepts discussed above. In comparison, the main dimensions in the other three samples were more general in nature. These and other differences between samples are discussed hereafter.

11.2 DIFFERENCES BETWEEN SAMPLE RESULTS

If there were few similarities between samples, in contrast there were many differences between various aspects of the results. These differences can be divided into three main categories: technical and methodological differences and differences in the nature of dimensions. These three categories are discussed in turn in the next two sections: technical and methodological differences are discussed first because it will be seen that they are useful in analysing and understanding the differences which were found in the nature of prominent dimensions extracted in the various samples.

11.2.1 Technical differences in results

In order to facilitate the discussion and understanding of the technical differences between the results of the four samples, some aspects of the results for each sample are summarized in Table 11.1. Firstly, the table lists the smallest percentage of variance explained by a first component in an individual grid in a sample. The second column presents the highest percentage of variance explained by a first component in an individual grid; the percentages of all the other first components of the sample fall in between these two figures. The "range" is the difference between the highest and the lowest percentages. The mean percentage of variance explained by the first component in all the grids of a sample can be found in the fourth column. The fifth set of figures is the mean number of cells in the grids of a sample, a cell being the rating of one element

TABLE 11.1

COMPARISONS OF SOME TECHNICAL ASPECTS OF SAMPLE RESULTS FROM INDIVIDUAL GRIDS

Sample	Smallest first root as %	Largest first root as %	Range	Mean of first root as %	Mean number of grid cells	Mean number of components (Bartlett)	Mean number of components (Kaiser)	Mean number of components with significant loadings	Mean number of clusters
GR	19.9	34.0	14.1	27.3	312	3.3	5.6	3.8	3.8
UG	22.1	36.4	14.3	29.4	1012	10.2	7.0	7.8	3.6
SA	19.6	87.9	68.3	41.2	420	7	5.2	4.9	3.3
PH	55.4	68.1	12.7	62.6	54	1	2	1.7	1.7

on one construct. Then, the mean number of components found significant by the Bartlett test and by Kaiser's criterion of a latent root of 1 are listed. In the eighth column, the figures indicate the mean number of components in which at least one significant loading (element or construct) was found. Finally, the mean number of clusters extracted from the grids of a sample is listed in the last column.

One of the most important features of this table is the difference in the percentage of variance explained by first components in the various samples. As far as these percentages are concerned, the first two samples (GR and UG) are relatively identical; both samples have a small range and an average between 25 and 30 percent. The SA sample has a higher average percentage, but the most noteworthy feature of this sample is that percentages range between 19.6 (the lowest figure of its column) and 87.9 (the highest figure of its column), for a range of 68.3 percent. As for the PH sample, its average percentage is the highest of the four, despite the sample range being the smallest of the four. It is argued in section 11.2.3 of this chapter that differences in the range of elements within each list explain some of these observed differences.

In turn, these results are important for the interpretation of the other results in Table 11.1. The more variance is accounted for by a first component, the less variance remains to be accounted for by other components, and therefore the less likely it is that other components emerge as being significant. Thus, a narrow range usually indicates that all grids in a sample yield more or less the same number of significant dimensions. In contrast, a large difference in percentages is an indication that there are important differences in the number of dimensions within the sample.

In the PH sample, whose range is the smallest, the number of significant components found by the Bartlett test was identical in all three grids, and so was the number of components retained by Kaiser's criterion. In contrast, in the SA sample whose range was the widest, the number of dimensions with significant loadings varied between one and seven.

The average percentage of variance can be an indication of the number of dimensions found significant. For instance in the PH sample, where the average is highest, never were there more than two significant components. However, in the GR sample, whose average was lowest, there were never less than two significant components.

PH grids consistently yielded a small number of components. GR and UG grids also showed a fair measure of agreement but on a larger number of components. It was in the SA sample that the most important differences were found between respondents in the assessment of identical hazards.

Table 11.1 shows almost identical smallest and largest percentages, ranges and mean percentages for GR and UG samples. On the basis of what has just been said, one would expect to find almost identical mean numbers of dimensions also. Yet, the mean numbers of components in both samples, particularly those for the Bartlett test, are quite different.

Another important observation about the figures in Table 11.1

must be pointed out at this stage. The figures suggest that the number of components in each sample is directly influenced by the size of the grid used in the various samples. There is a perfect rank-order correlation, for instance, between grid size (mean number of grid cells) and mean number of components as determined by the Bartlett test; the correlation is also equal to 1.0 between grid size and mean number of components including at least one significant loading. Furthermore, PH grids (the smallest of the four types of grids) accounted for the smallest mean number of components by Kaiser's criterion; in contrast, UG grids (the largest) yielded on average the largest number of components with a latent root equal to or greater than 1. It must be pointed out, however, that each correlation was calculated only of four pairs of figures. The possibility that variables other than grid size influenced the number of significant components cannot be excluded. A number of these variables are examined later in this chapter.

Thus, it appears that the larger a grid is, the more components it yields. The reason for this is quite simple. The larger a grid is, the more ratings it encompasses. The more ratings there are, the greater the total variance of the grid is. Even if unique variance is not taken into consideration, it remains that larger grids generate more common variance. As there is more common variance to account for in larger grids, it is very likely that more components can be identified which explain this variance. A small portion of a large amount of variance may sometimes be more sizeable than even a large portion of a very small amount of variance.

It is noticeable, however, that grid size appears to have had little or no influence upon the number of clusters. This can most probably be explained by the fact that what is analyzed in identifying clusters is the covariance between two constructs rather than the total covariance of a grid. The number of constructs in a grid is not likely to have much effect on the correlation between any two of them.

It can be argued that grid size can nevertheless have some influence on the number of clusters. For instance in a small grid the minimum significant correlation (the cut-off point) was very high and thus limited the possibility of generating many clusters; high correlations were usually scarce even in large grids. However, it can also be argued that the lower the cut-off points were, the larger the clusters were and therefore the fewer constructs were left to form subsequent clusters. Thus, because of these two opposite effects, it was not likely that cut-off points (and indirectly, grid size) would have a well-defined influence upon the number of clusters. This is substantiated by the fact that three samples (GR, UG and SA) all had mean numbers of clusters between three and four.

Cut-off points (determined by the size of the grid) probably influenced the number of components. High minimum loadings in small grids may well have impeded the emergence of many components. In large grids, however, low cut-off points may have made it easier for elements and constructs to be identified as significant members of components. Unlike cluster analysis, in principal component analysis a construct could be found in more than one component. Thus, the influence of cut-off points on principal component analysis was similar to the influence of the amount of common variance. Larger grids, because they involved more common variance and because they were subjected to lower cut-off points, tended to yield more components. But, as pointed out earlier, grid size had little effect upon the number of clusters.

It is interesting to note, however, that the number of labelled dimensions was generally closer to the number of clusters than to the number of components. For instance, in the UG sample, the mean number of dimensions to which a label (even if it was a very doubtful one) was affixed was 4.5; this is much closer to the mean number of clusters than to any of the mean numbers of components in Table 11.1. Similarly, in the SA sample, the mean number of labelled dimensions was 3.9; again, this is closer to the mean number of clusters than to the mean numbers of components. These findings seem to confirm the point made in previous chapters that cluster analysis was a more stringent (and possibly a more accurate) technique than principal component analysis.

Discrepancies between the numbers of labelled dimensions and the numbers of components found significant by various techniques can be explained in two ways. Firstly, as pointed out in Chapters 8 and 9, there are reasons to believe that the Bartlett test, performed on as many components as there were constructs in a grid, tended to overestimate the number of significant components in larger grids. That grid size may have introduced bias in the test is illustrated by discrepancies (between the number of components retained by the test and the number of clusters) increasing as the size of the grids increases. Secondly, although many components accounted for at least one significant loading, it often happened that the contents of a component were scarce and difficult to interpret.

As could be expected, those components involved in the discrepancies were minor components, usually explaining very little common variance. What is important for the purpose of this research is that the main dimensions were generally robust. It was pointed out in Chapter 5 that robust dimensions were the dimensions identified unanimously both by cluster analysis and by principal component analysis. Therefore, although grid size influenced the number of components found significant, this influence did not much alter the most important observations of this research. The main causes of the discrepancies between the numbers of components and the number of interpretable or robust dimensions appear to be flaws or inconsistencies in the various methods for deciding on the number of significant components.

It does not seem that any individual characteristic of respondents in the various samples would explain the differences in the number of robust dimensions between the samples. The youngest sample (UG) had the largest mean number of robust dimensions, but the sample with the lowest mean number (PH) also included young respondents. The two samples which included only men showed the highest (GR) and the lowest (PH) mean number of clusters. Other variables, such as number of respondents in a sample, proportion of respondents with a university education or whose first language is not English, were looked into but none was found to be in direct relation to the numbers of dimensions.

There are some indications that knowledge and experience of health and safety matters might explain some differences in the number of dimensions. For instance, all GR respondents were members of the health and safety committee in their workplace, and all UG respondents were receiving university tuition in health and safety. It is interesting to note that these two samples show the two highest mean numbers of clusters. However, these numbers are not much larger than the corresponding figure for the SA sample. Furthermore, the SA sample tended to show higher mean numbers of components than did the GR sample. Finally, as will be pointed out later, it was in the SA sample that the greatest diversity between individual results was observed.

11.2.2 Methodological differences

A number of aspects of the grid technique were also looked into. It will be remembered that the grid techniques used in the four samples differed in a number of ways: the way in which elements were elicited, whose time (respondents' own or employers') was taken up for element elicitation, the way in which constructs were elicited, whose time was taken up for construct elicitation, the type of rating scale used, and whose time was taken up to score the grids. It was pointed out in previous chapters that some of these variables had an influence on the ratings (e.g. the type of rating scale influenced the distribution of ratings). However, none of these variables seemed to be related to the numbers of robust dimensions.

Since neither differences in individual characteristics of respondents nor differences in grid techniques could explain the variations in the numbers of dimensions between samples, it would appear that only differences in the actual contents of the grids could explain these variations. This is what is argued in the next section of this chapter, where the contents of the grids and of the dimensions in the various samples are reviewed and compared.

Other technical aspects of the results also appear to indicate that variations between samples of the main dimensions were explained mainly by variations in the contents of the grids. For instance, it was mentioned in section 11.2.1 that the difference within a sample between the highest and the lowest percentages of common variance explained by a first component was an indication of the diversity of respondents' thinking about similar hazards. In other words, a wide range of percentages suggests important differences in individual hazard assessments. Thus, explanations were needed at two levels. Firstly, reasons had to be sought for differences between respondents within a sample. Secondly, it was necessary to understand why some samples showed differences between respondents greater than in other samples.

Three analyses were performed in an attempt to explain differences between individual results. Firstly, in the GR sample a Mann-Whitney U-test was performed on percentages of variance explained by first dimensions. Respondents were sorted into those involved and those not involved in the production process. No statistically significant difference was found between the percentages for the two groups of respondents. Secondly, UG respondents were sorted into British-born and foreign-born respondents, and their percentages of variance were subjected to a Mann-Whitney U-test. Again, no significant difference was found. Finally, the percentages of variance were compared for male and female UG respondents on the same statistical test, which did not reveal any significant difference.

It must be pointed out that the analyses described above were performed within samples which showed comparatively small ranges, i.e. in which differences between respondents were small. It seems possible that some individual characteristics (e.g. demographic variables, age, etc) could explain individual differences between results. However, the size of these differences makes them difficult to explain. Furthermore, no other variable was controlled within any sample.

Therefore, comparisons between samples were turned to. It seemed possible that a variable which could explain why individual differences were greater in some samples than in other samples might provide indications about the reasons for individual differences. All differences in respondents' individual characteristics and in grid techniques between samples as described earlier were looked into in an attempt to explain range differences between samples. No conclusive evidence could be found. Only differences in grid contents remained to be explored in search of an explanation for the fact that the range of differences within the SA sample was greater than the ranges of the other three samples put together.

11.2.3 Differences in grid contents

As pointed out in the previous section, there are indications that sample differences in hazard assessment patterns were generated by differences in the actual contents of the grids. This section is an examination of further evidence of relations between grid contents and hazard assessment patterns. This examination is followed by a discussion of the reasons why grid contents influenced the identification of important dimensions.

In the previous section, the influence of grid contents on emerging dimensions was inferred from the fact that no other variable seemed to account for sample differences in variance ranges and numbers of dimensions. However, there is more direct evidence of the relationship between grid contents and emerging dimensions.

For instance, Table 11.2 presents the number of different dimensions identified in each sample. The table also shows the total number of different constructs used by respondents in each sample. In three samples (GR, UG and PH) all respondents used identical constructs. In the SA sample, no two grids contained exactly the same list of constructs. As pointed out in Chapter 10, however, some constructs were found more than once in some grids and others were found in more than one grid. The figure in Table 11.2 is the number of <u>different</u> constructs found throughout the SA sample. Finally, the table lists the number of elements rated in each sample.

Perhaps the most noticeable feature of Table 11.2 is the fact that the number of different labels used in each sample is directly proportional to the number of constructs, there being approximately three times as many constructs as there were labels in each sample. The rank-order correlation between the two sets of figures is 1.0, and the parametric correlation in 0.996; both correlations are highly significant ($p < .05$ and $p < .01$). This observation seems to indicate that the diversity of dimensions found

TABLE 11.2

NUMBER OF DIMENSIONS, OF CONSTRUCTS AND OF ELEMENTS IN EACH SAMPLE

Sample	Number of dimensions identified	Total number of different constructs	Number of elements
GR	4	17	19
UG	9	27	38
SA	23	64	21
PH	2	6	9

in a sample was a direct function of the diversity of constructs used in that sample. These results are in line with the observations that grid size influenced the number of significant components. However, while grid size could not account for differences in the numbers of labelled dimensions, it seems that the number of constructs can.

Table 11.2 contains an indication of the reasons why numbers of constructs are different. For three samples, the number of constructs is proportional to the number of elements, there being a few more elements than constructs. It seems plausible that the number of constructs should be proportional to the number of elements, since constructs are elicited from elements. The more elements there are, the more triads it is possible to form and thus the more possibilities there are of eliciting constructs. In the UG and PH samples, it seems that every three elements produced two constructs on average. The ratio is different in the GR sample because only those constructs most frequently elicited in the pilot experiments were used in the final version of the grid. The one noticeable difference is found in the SA sample, where there are three times as many constructs as there are elements. Therefore, the number of elements cannot be the only explanation of the number of constructs.

One plausible, and maybe the most likely, explanation appears to be the difference in the diversity within the four lists of elements. In other words, there are indications that in three samples (GR, UG and PH) the list of elements was relatively homogeneous whereas the fourth list (in the SA sample) contained fairly heterogeneous elements. All GR elements were occupational hazards. All the elements rated in UG grids had at least one characteristic in common - that they were home hazards. There was at least one characteristic common to all PH elements: they were hazards inherent to spelaeology. Apart from the fact that all of them were hazards, there was practically nothing in common between all SA elements. Furthermore, there were 38 UG elements all of which were home hazards; and yet, "accident in the home" was but one of the 21 elements in SA grids.

There were other indications of the homogeneity of some series of elements and of the heterogeneity of SA elements. For instance, 77% (176 out of 228) of all ratings on GR construct 12 were 1's; in other words, a majority of GR respondents rated a majority of elements as "preventable". In contrast, three SA respondents considered the concept of preventability as significantly useful for sorting out the elements.

Thus, there are indications that for three samples the number of constructs (and consequently the number of labelled dimensions) were a direct function of the number of elements. In the SA sample, the number of dimensions was also a function of the number of constructs; but is is very likely that the large number of constructs was caused by the diversity among the list of elements.

There is another indication that the diversity of SA elements led to a diversity of hazard assessment patterns within the SA sample. It was pointed out earlier that the range of percentages of variance explained by first components was much larger in the SA sample than in the other three samples. It was argued that such a large difference indicated that hazard assessment patterns varied greatly between respondents within the sample. Furthermore, it was argued in Chapter 10 that, unlike in other samples, no major dimension emerged as common to most grids; it was pointed out that large dimensions in other samples were fragmented by SA respondents. Table 11.3 presents additional indications of hazard assessment patterns more diverse in the SA sample than in the other samples. The table lists the smallest and the largest number of dimensions found in individual grids of each sample for four indicators: the number of components found significant by the Bartlett test, the number of components found significant by Kaiser's criterion, the number of components including at least one significant loading, and the number of clusters. Ratios of the smallest to the largest number are also listed in each case. These ratios have the advantages of being a common basis of grid size and of removing part of the bias introduced by the occasional unreliability of the various methods of identifying significant dimensions. A ratio of 1:1 indicates that all grids in a sample yielded the same number of dimensions; a ratio of 1:14 reveals that one grid in the sample yielded 14 times as many dimensions as the grid which yielded the fewest dimensions.

It is interesting to note, in Table 11.3, that the four ratios of the SA sample are consistently the highest of their category. This means that the disparity in the number of dimensions was proportionally greater in the SA sample than in any other sample, irrespective of the method used for deciding on the number of significant dimensions. It is also noticeable in Table 11.3 that whereas the ratios for the other three samples appear to be within the same order of magnitude the SA ratios tend to stand out on their own as if in another order of magnitude.

TABLE 11.3

RATIOS OF THE SMALLEST TO THE LARGEST NUMBER OF DIMENSIONS IN EACH SAMPLE

	Number of components retained by the Bartlett test			Number of components retained by Kaiser's criterion			Number of components with significant loadings			Number of clusters		
Sample	Least	Most	Ratio	Least	Most	Ratio	Least	Most	Ratio	Least	Most	Ratio
GR	1	6	1:6	5	7	1:1.4	3	6	1:2	2	5	1:2.5
UG	6	25	1:4.2	6	8	1:1.3	6	10	1:1.7	2	5	1:2.5
SA	1	14	1:14	1	8	1:8	1	8	1:8	1	6	1:6
PH	1	1	1:1	2	2	1:1	1	2	1:2	1	2	1:2

A number of variables were looked into in order to find an explanation for these differences of ratios. However, there is only one characteristic which the GR, UG and PH samples have in common and which is different in the SA sample. It appears that the heterogeneity of the elements in the SA sample occasioned more diverse, and sometimes more differentiated, hazard assessment patterns. In contrast, it is probably because their elements were fairly homogeneous that GR, UG and PH ratios were quite similar.

It was argued earlier that the more diversity there was within a series of elements, the more characteristics or constructs there were which could be used to assess or to sort the elements. This does not mean, however, that all respondents use the whole array of constructs in their assessment. Kelly's (1955) fourth corollary states that for a specific category of elements a person's construing rests upon a finite number of constructs. Kelly also quotes research results which suggest that for a given category of elements, different people use construction systems whose sizes are similar, i.e. whose number of constructs are comparable. It was pointed out in Chapter 10 that in the SA sample, the only sample in which respondents used truly individual construction systems, the numbers of constructs varied between 15 and 25. In contrast, the total number of different constructs in the SA sample was 64.

It seems plausible that, given an average number of 20 constructs in each SA grid, the greater the number of constructs which can be used, the greater the number of possible sub-sets of 20 or so, and thus the greater the probability that individual sub-sets are different in whole or in part. Since diverse constructs appear to be responsible for differences in the numbers of dimensions, it seems possible that diversity in the elements were indirectly responsible for inter-respondent variability in hazard assessment patterns.

Such an explanation rests upon the assumption that the diversity within the series of elements accounts for the diversity of constructs. This assumption was examined by looking into various other characteristics of the elements rated by the four samples. None of these other characteristics could be found to account for the differences in the numbers of constructs. For instance, it can be argued that some of the UG and PH elements resulted from other hazards. However, such an explanation could not account for the fact that GR grids in which none of the elements were consequences of other hazards to the same extent as in the other two samples yielded dimensions relatively similar to those from UG and PH grids. The fact that some SA elements were worded as accidents rather than as hazards was also discarded as an explanation on the basis that Golant and Burton's (1969) elements were also worded as accidents and that these yielded dimensions more comparable to GR, UG and PH dimensions than to SA dimensions.

Other alternative explanations were also sought for the diversity of constructs. All the differences in individual characteristics and in grid techniques mentioned in the previous section of this chapter could be discarded as having little influence on the number of constructs in each sample. For instance, the numbers of respondents in the GR and UG samples were almost identical but UG respondents were university students just as SA respondents were, yet the latter generated nearly 2.5 times more constructs.

One of the important differences in grid techniques was the degree of "individuality" of the grids, in other words, the degree of respondents' participation in the elaboration of their own grid. It will be remembered that PH and UG respondents actively participated in the elicitation of the elements as well as of the constructs of the grids they used. And yet, despite group discussions in both the UG and the PH samples and the possible cross-fertilisation that group discussion may involve, there were more constructs in the SA sample than in the other three samples together. It can be argued that the diversity of constructs in the SA sample is attributable to the fact that construct elicitation was carried out individually with SA respondents. However, this argument would hardly explain why the PH and UG samples, whose element and construct elicitation techniques were very similar, had such different numbers of constructs. Thus, differences in individual characteristics and in grid techniques could not explain the differences between the numbers of constructs.

There may have been other variables, not controlled in this research, which could explain why diversified dimensions were identified in the SA sample. Nevertheless, many indications converge to account for the fact that the three samples in which relatively homogeneous lists of elements were used yielded dimensions which were more coherent than those identified in grids where elements were quite heterogenous. These similarities and differences in hazard assessment patterns are reviewed in the next sections of this chapter.

11.3 A REVIEW OF THE MAIN CONCEPTS

When the results of a sample were analyzed, emerging dimensions were discussed separately. Very few references were made to similarly labelled dimensions in other samples. The possible relations between dimensions, even those emerging from the same grid, were seldom discussed. And yet in comparing the main concepts emerging from the four samples, some interesting observations come to light. The next sections discuss in turn the various dimension labels used throughout the analyses of the results of the four samples in order to gain further insights into the main concepts underlying hazard assessment.

Some discussion of the labelling process is needed at this stage in order to understand some of the points raised in the sections hereafter. The label of a component was generally chosen on the basis of the construct, amongst the most prominent ones, which best explained why the significant constructs of the component should gather the way they do. The label was verified by checking whether the characteristic underlying the label was the best possible way of explaining why elements were sorted into positive and negative ones. Clusters were labelled by identifying the characteristic which their constructs had in common. Then, as the results of cluster analysis and of principal component analysis were usually in agreement, the corresponding cluster and component were compared to make sure that the labels of each were apposite.

Occasionally, none of the constructs in a dimension seemed to reflect accurately the main theme of the dimension. This tended to occur in large dimensions encompassing many constructs. These constructs could usually be divided into two or three categories. The label was then chosen so as to reflect the characteristic which might explain the gathering of these categories into one dimension.

There is a certain amount of subjectivity involved in the labelling process. Comparing corresponding clusters and components removed some of the subjectivity. More uncertainty about the labels is removed by the fact that similarly labelled dimensions in different samples showed some constants. These constants are outlined in the next sections.

11.3.1 Scope of human intervention

The label "scope of human intervention" was used in three samples: GR, UG and PH. In all three cases, it was mentioned that the dimension thus labelled had more than one facet. When comparisons are made between the three samples, a remarkable degree of similarity between the facets can be seen.

For instance, it was pointed out in Chapter 8 that the first GR dimension contrasted hazards which were the results of an operator's action(s) with hazards which were part of the environment or of the work system. In Chapter 9, the UG dimension was described as opposing built-in features of the environment to hazards attributable to the person concerned. In the GR dimension, hazards which were attributable to an operator were seen also as correctable by the operator and as preventable. It was mentionned in Chapter 9 that in the UG dimension those hazards seen as attributable to a person were also construed as being under the person's control. The most prominent PH dimension is very much along those lines; hazards attributable to a person ("lack of thought", "due to my party") were assessed as "controllable" whereas hazards not directly attributable to someone ("bad luck", "due to another party") were deemed "uncontrollable".

There were no constructs in GR or UG grids whose wording was identical to the PH constructs in this dimension. There was however one construct similarity between PH and UG grids. PH construct number 5 was worded "poor planning/poor moving"; UG construct number 12 read "this is the result of poor planning". In both samples this construct occasionnally belonged to the "scope of human intervention" dimension without being altogether a prominent or a frequent member of this dimension. On those occasions when the construct belonged to this dimension hazards construed as attributable to someone always had a connotation of "poor planning".

There were further construct similarities between GR and UG grids. For instance, in the GR dimension, hazards attributable to someone were seen as having "nothing to do with design"; UG respondents disagreed with the statement that these hazards were "due to bad design". In GR grids "anyone can put it right" was used to describe these hazards (the opposite pole of the construct being "takes a specialist to put it right"); similarly in UG grids, hazards attributable to someone were associated with disagreement with the statement "it takes a specialist to put this right" and were rated as "very easy to put right". Furthermore, it is interesting to note that, in both samples, hazards seen as being under a person's control were rated as "temporary".

It must be noted that in all three samples with very few exceptions this dimension systematically emerged as the first to be extracted. It was also consistently among the most robust within the various grids.

Because of everything that has been said so far about this dimension, it was somewhat surprising not to find any dimension thus labelled in the SA sample. What was found instead in this sample was a series of more specific dimensions focused on one or more of the facets of the concept of "scope of human intervention".

For instance, the preoccupation about the attributability of a hazard to the person facing the danger or to someone or something else ("bad luck", "other party", built-in design feature) is found as a separate and specific dimension in SA grids. A dimension labelled "natural/man-made" was found in six grids. As pointed out in Chapter 10, these dimensions included constructs such as "human cause/no human cause", "self responsible/self not responsible", "blame assignable/no blame assignable" and, of course, "natural/man-made". The dimension labelled "attributability of blame", found in another grid, had similar contents.

The concepts of a person's control over hazards, prominent in "scope of human intervention", was found in distinct dimensions labelled "controllability". Five SA grids yielded such a dimension. The contents were focused on constructs such as "own control/out of own control", "self responsible/self not responsible" and "rely on others/rely on self".

The dimension labelled "preventability", found in three SA grids, also reflected another important facet of "scope of human intervention" as defined in the GR and UG samples. Its main constructs were: "preventable/unpreventable", "avoidable/unavoidable", and "precautions/no precautions".

There were indications that these four dimensions were interrelated in the minds of SA respondents. For instance, in some grids, some of the constructs mentioned above loaded significantly on two of these dimensions. There were also cases where a few constructs of one dimension in one grid were found with the same polarity in another of these dimensions in another grid.

There were two other dimensions which showed the same signs of relatedness between themselves and with the four dimensions mentioned above. Firstly, a dimension in one grid received the label "Easy/difficult escape"; its main constructs were "easy to escape/ difficult to escape", "own control/out of own control" and "rely on others/rely on self". The possibility of dodging the potential consequences was an inherent facet of this dimension; therefore, the fact that this dimension was related to the previous four is interesting in view of the fact that GR construct 8 ("Easy to avoid consequences of danger/impossible to avoid consequences of danger") was one of the prominent members of the "scope of human intervention" dimension.

Secondly, the most frequently identified dimension of the SA sample, i.e. "Voluntariness of activity", also showed some relatedness to other facets of the most prominent GR, UG and PH dimension. The main constructs of this SA dimension were "necessary/ unnecessary activity" and "occupational/not occupational". But other constructs from the SA dimensions referred to above ("natural/ man-made", "human cause/no human cause" and "own control/out of own control") were also members of this dimension in a few of the eight grids where it was found. Two other SA dimensions ("occupational/not

not occupational" and "acceptance of risk") seemed to be related to "voluntariness of activity" but not to the other facets of "scope of human intervention". Despite the relatedness of "voluntariness of activity" to the four SA dimensions referred to earlier, the underlying concept of control over one's exposure did not emerge as a significant facet of the main dimension in the other three samples.

It seems that "scope of human intervention" was a fairly general dimension which SA respondents fragmented into various more specific dimensions. Since the general dimension in the GR, UG and PH samples were as robust as the specific dimensions in the SA sample, it would appear that it is in fact the diversity of SA elements which led to the type of differentiated thinking found in the SA sample.

In three samples "scope of human intervention" was prominent over "severity of consequences". In the SA sample, however, there are no clear indication of such prominence. This point is discussed hereafter.

11.3.2 Severity of consequences

The severity of potential consequences was the second most important concept underlying hazard assessment in three (GR, UG and PH) samples. Thus, the second dimension extracted from both the $\overline{X}$ and the $\overline{Z}$ grids in the GR sample was labelled "dreadfulness"; in retrospect, the label "severity of consequences" seems more apposite since this dimension is concerned with the consequences themselves rather than with the fear they may elicit. In the UG sample, the label "dangerousness" was affixed to a dimension in ten of the 11 individual grids as well as in both the $\overline{X}$ and the $\overline{Z}$ grids. This dimension was extracted in second place in seven individual grids and in both mean grids, in first place in two individual grids and in third place once. The label "dangerousness" was chosen on the basis that the dimension encompassed the concept of likelihood of occurrence as well as the concept of severity of consequences. In two of the three individual PH grids as well as in the $\overline{X}$ grid, indications of a significant second dimension could be found; this second dimension was tentatively labelled "severity of consequences". It must be pointed out, however, that a second dimension was not labelled in the same way in the third individual PH grid.

The second dimension in the GR, UG and PH samples and the SA dimensions which reflected the various facets of the concept of severity of consequences showed many similarities. Firstly, GR construct number 10 ("very likely to kill/very unlikely to kill") and PH construct number 6 ("likely to kill/unlikely to kill") were both the most prominent in their respective dimension; UG construct 27 ("how likely/unlikely is this to cause death") was among the first constructs of its dimension. Secondly, GR construct 9, which read "permanent disability/only a trivial injury", and UG construct 21, worded "how serious/trivial an injury could this cause", were both in third position in their respective dimension. Thirdly, again in this dimension, GR construct 13 ("takes a specialist to put it right/ anyone can put it right") and UG constructs 20 ("how easy/difficult is it to put this right") and 10 ("it takes a specialist to put this right") were members of the second dimension. Finally, other prominent constructs in this dimension were "never encountered in my job/very often encountered in my job" (GR number 1) and "one comes across this..." ("frequently/infrequently"; UG number 26).

The second UG dimension appears to be more general in nature than the corresponding dimension in the GR and PH samples. The former encompassed concepts such as direct/indirect hazard and likelihood of occurrence which were not present in the other two samples.

In contrast, again SA dimensions were more specific. It must be pointed out that there was a "severity of consequences" dimension in seven individual SA grids. But, throughout the 17 SA grids a number of dimensions were concerned with various facets of the concept of "severity of consequences".

Firstly, a dimension labelled "nature of consequences" was found in one grid. In this dimension, hazards were sorted out into those which caused mainly material damage and those whose most important consequences were losses of lives. In another grid, a dimension received the label "extent of material damage"; as the name implies this dimension was concerned with the amount of destruction (other than loss human lives) which hazards could cause.

Three dimensions were concerned with the potential impact of hazards on people. The first of these dimensions, found in five grids, differentiated between hazards affecting a small area (an individual, a small geographic location) and hazards affecting a large geographical area and thus also affecting a large number of people; that is why the label "widespread/localized" was chosen for this dimension. The second dimension was labelled "size of population at risk" and was found in one grid. The main theme of this dimension was much the same as that of the previous dimension except for the geographical connotation of the latter. The third dimension was identified five times and labelled "number affected". Whereas the previous dimension assessed whether everybody or only special populations were at risk, this dimension assessed the number of people affected at once by a hazard. The label "number affected" was chosen on the basis that the dimension was concerned with material losses as well as injuries, lives lost and other impacts (e.g. bereavement) on the persons concerned.

Finally, another dimension was extracted from one grid and received the label "short-term/long-term effects". This dimension was concerned with the delay between the occurrence of a hazard and the appearance of the effects.

It was pointed out earlier that the second UG dimension was labelled "dangerousness" on the basis that it encompassed the concept of likelihood of occurrence. This concept emerged as a distinct dimension, labelled "likely/unlikely", in three SA grids.

Most of these facets were encompassed together in the general dimensions as identified in the GR, UG and PH samples. Whilst they emerged as distinct dimensions in the SA sample, there were indications of inter-relatedness between them in that sample. For instance, the construct "large/small consequences" was the dominant feature of the dimension labelled "severity of consequences"; the same construct was also among the most significant members of many of the dimensions mentionned above. The construct "major/minor" was also common to many of these dimensions. The construct "personal/ impersonal", a key concept in the "number affected" dimensions, was also found in "severity of consequences" and "widespread/localized" dimensions.

The concept of likelihood of occurrence, encompassed in the UG dimension labelled "dangerousness", showed some affinity with the various facets of "severity of consequences" found in the SA sample. For example, the construct "high/low risk of accident", the focal point of the three "likely/unlikely" dimensions, was also found in "widespread/localized", "short-term/long-term effects", "number affected" and "severity of consequences" dimensions.

Thus, it seems that the assessment of the severity of potential consequences was carried out with varying degrees of complexity in the four samples. In the UG sample "severity of consequences" was one of the two main concepts encompassed in the label "dangerousness". In the GR and PH samples, "severity of consequences" emerged as a distinct concept. The dimensions thus labelled in both samples had some characteristics in common. However, the PH dimension was very small and sometimes difficult to extract. These difficulties, inherent in small grids, create a risk of overinterpretation of the dimension.

In the SA sample, in addition to dimensions labelled "severity of consequences", six other types of dimensions were found to center around concepts associated in either the UG or the GR dimension. The concept of likelihood of occurrence, associated with the concept of severity in the UG dimension, emerged as a distinct dimension on three occasions in the SA sample.

One important difference between the SA sample and the other three samples was in the position in which the "severity of consequences" dimensions were extracted. The "dangerousness" dimension in the UG sample and the "severity of consequences" in the GR and PH samples were quite systematically extracted in second position. In contrast, five of the seven "severity of consequences" dimensions in the SA sample were the first (and on one occasion the only) dimension to be extracted from their respective grids. The labels "widespread/localized" and "number affected" were affixed to the first dimension in two grids each. The concept of "riskiness", akin to the "dangerousness" dimension in the UG sample, was the main theme of a first dimension in one SA grid. In contrast, one or another of the facets of "scope of human intervention" emerged as prominent only in five SA grids.

These indications may suggest that in the SA sample "severity of consequences" and the various concepts it encompasses may have been prominent over the various facets of "scope of human intervention". It can be argued that some SA elements depict disasters whose consequences are of an order of magnitude different from that of consequences from elements in the other three samples. However, items in Golant and Burton's (1969) and in Fischhoff et al's (1978) research also depicted major hazards some of which were similar to SA elements (e.g. "nuclear power" and "accidental release of nuclear radiation"). And yet, in both studies, dimensions reflecting "severity of consequences" were extracted after dimensions related to the notion of control. In any case, in the SA sample, indications of the possible prominence of "severity of consequences" over "scope of human intervention" are not convincing. Further research is needed on this point. Some suggestions are provided in the next chapter.

In all four samples, there were indications that "scope of human intervention" and "severity of consequences" were correlated. The inter-relatedness of the various concepts is discussed later on in this chapter. Before this, however, other concepts which have been identified as important in the various samples deserve some consideration.

11.3.3 Dread

It has been hypothesized that the second most important dimension throughout the four samples would be labelled "dreadfulness". It turned out, however, that "severity of consequences" was a more apposite label in two samples and that "dangerousness" was more appropriate in the UG sample. As for the SA sample, "severity of consequences" and other more specific labels were used.

In the last three chapters, it was argued that these labels were each more appropriate than "dreadfulness" on the basis that the notion of dread as such was not present within these dimensions. It is interesting to note, however, that three SA dimensions bearing different labels had the concept of repulsiveness of hazards in common.

In grid SA01, the unique dimension was labelled "hideousness" on the basis that constructs having "hideous" at one pole were among the most significant members of the dimension. In grid SA02, the "ugliness" dimension could be said to be akin to the other three dimensions discussed in this paragraph. However, this dimension appears to be an evaluative rather than an emotional pattern. In grid SA04, constructs referring to "severity of consequences" formed the core of the cluster and had the highest loadings. In grid SA08, the prominent constructs tended to refer to the concept of horror, which influenced the choice of the label.

The unique dimensions in grids SA01, SA04 and SA08 had some similarities which are not obvious from the labels. For instance, the SA01 and SA09 dimensions both involved secondary constructs referring to "severity of consequences". All three dimensions encompassed "dangerous/not dangerous" constructs. SA04 and SA08 dimensions both included constructs related to fright and to pain, in addition to constructs having the adjective "unpleasant" at one pole.

Although none of the three respondents referred to "dreadfulness", and despite the different labels used, a strong undertone of dread appears to be present in the SA01, SA04 and SA08 dimensions. Because of the various connotations of these three dimensions, they seem to be the only emotional dimensions amidst a spate of evaluative dimensions in the SA sample. Thus, it would seem that "severity of consequences" is an evaluative reaction to hazards whereas "dreadfulness" is an emotional reaction.

It seems likely that there were evaluative dimensions among the minor, non-significant dimensions of the three grids. Assuming that there were, this would mean that an emotional reaction towards hazards would take precedence over a detached assessment of hazards. One is reminded of a panic reaction. But further research would be needed on this point.

It is somewhat surprising that no emotional dimension was identified in the other three samples. An additional look at the constructs used by these samples reveals an absence of constructs having an emotional undertone. It seems likely that the absence of emotional dimensions was caused by the absence of emotional constructs.

Some suggestions can be put forward to explain the absence of emotional constructs in the GR, UG and PH samples. For instance, it may be that only SA elements are large enough hazards to be susceptible to provoke emotional reactions or panic. GR respondents did not use their own constructs. Had they done so, emotional constructs might have appeared in GR grids. As for UG and PH respondents, it must be remembered that they mentioned their constructs during group discussions. It is possible then that UG and PH respondents who might have mentioned emotional constructs were reluctant to do so in the presence of other participants. Again further research would be needed in order to shed some light on these arguments.

11.3.4 Familiarity, prominence, obviousness

When discussing GR dimension number 3 Champion (1977) affixed to it the label "familiarity". It was argued in Chapter 8 that a label such as "prominence" or "obviousness" might be more appropriate. Some support for Champion's label is found in similar dimensions in other samples.

Some dimensions labelled "familiarity" and others labelled "prominence" were found in individual UG grids. Familiar hazards were rated as often encountered, quite obvious and not conditional upon a chain of events. Because they were familiar, however, these hazards were seen as unlikely to cause an accident. In contrast, prominent hazards were rated as quite obvious, as hazards in themselves and not dependent upon a chain of events: such hazards were also assessed as being very infrequently encountered. Although the contents of the two "familiarity" dimensions in SA grids were limited, they seem to fall in line with the contents of the corresponding UG dimensions.

In GR dimension number 3 the construct poles labelled "very often encountered in my job" and "easily spotted danger" were associated at the same pole of the dimension. This is similar to the contents of the "familiarity" dimension in other samples. Thus, Champion's label for this dimension appears to be more appropriate than the two labels which were suggested.

In all three samples, the concepts of "familiarity" and "prominence" were not important aspects of hazard assessment. In the GR sample, the "familiarity" dimension was extracted in third place. In the UG sample, the four corresponding dimensions were extracted in third, fourth, fifth and sixth place respectively. These dimensions were fourth and sixth once each in SA grids. As for the "prominence" dimensions in the UG sample, one was found in third place, three were extracted in sixth place and one occupied seventh position.

11.3.5 Immediacy of danger

The last type of dimension found in more than one sample was the one labelled "immediacy of danger". This label was affixed to the fourth GR dimension and to the second dimension in individual grid UG01. Both dimensions had two constructs in common. In both cases "immediate dangers" were rated as "very likely to cause an accident" and as not being "dependent upon other things". However, the UG dimension also encompassed various aspects of "severity of consequences" not included in the GR dimension.

Thus the label "dangerousness" appears to be more apposite for the UG dimension. It was argued earlier that such a label had a twin connotation: "severity of consequences" and "likelihood of occurrence". Both connotations are present in the UG dimension. Furthermore, UG construct 25 ("How dangerous is this?") was the second most important construct in this dimension.

As for the GR dimension, it can be noticed that the two constructs mentioned above are very similar to the main contents of the "prominence" dimensions discussed in the previous section of this chapter. Thus, it seems appropriate to change the label of this fourth GR dimension to that of "prominence". Therefore, although the similarity of labels reflected a certain similarity of contents, identical labels were more probably the result of slight inadequacies in the original labelling process.

11.3.6 Other dimensions

There were a total of three types of UG dimensions and five types of SA dimensions found only in their respective samples. These eight labels are discussed briefly in this section.

Of these eight dimensions, only the "planning" dimension found in the UG sample has some claim to prominence. It was found in seven individual grids, and it was extracted as the third consensus dimension from the mean grids. It was pointed out in Chapter 9 that the concept of "planning" had two main connotations. Firstly, the concept refers to the amount of forethought that went into the preparation of one's actions. Secondly, the amount of care with which the lay-out of an environment was designed is also a prominent aspect of "planning". There were indications that the extent to which a hazard was a planning mistake was in turn an important facet of "scope of human intervention" as it emerged in the UG sample. "Planning" emerged as a specific dimension even in grids where a "scope of human intervention" dimension was found, much like specific facets of severe consequences emerged in those SA grids which had already yield a dimension labelled "severity of consequences".

"Frequency of occurrence" was a dimension found only in four individual UG grids. There are indications that frequent hazards are assessed as likely to cause an accident. The majority of SA constructs labelled "frequent/infrequent" were found in dimensions labelled "likely/unlikely". In addition, most constructs of the four UG dimensions were also found in the "dangerousness" dimensions of their respective grids.

As for the "origin of danger" dimensions, they were found in seven individual UG grids. There is also a possibility that the label could be apposite for the fifth UG consensus dimension which remained unlabelled in Chapter 9. These dimensions assess the extent to which the presence of a hazard is caused by the actions of the person involved in the situation. On the basis of common constructs, it seems that "origin of danger" is a specific facet of "scope of human intervention".

Among the five SA dimensions not replicated in other samples, one seems to have absolutely no link with any concept in other samples. It was found in two grids and was labelled "modern/traditional". This dimension appears to reflect the diversity of SA elements.

The same can be said of two other types of dimensions found only in SA grids. The label "enclosed/open-air" was used in three grids. The "occupational/not occupational" dimension was extracted from only one grid. To a certain extent, it is understandable that these concepts were not relevant in other samples. Thus, all GR elements were "occupational" hazards, whereas none of the UG and PH elements were. All PH elements were underground hazards, and all GR and UG elements were indoor hazards; thus, il all three samples all hazards were "encolsed". It is not certain, however, that the "modern/traditional" dimension was irrelevant in the GR, UG and PH samples.

The last dimension found only in the SA sample was labelled "moving/stationary", and was extracted from two grids. There was a construct similarly labelled in the GR grid. This construct was a significant member of the "severity of consequences" dimension. In the SA sample, "moving dangers" tended to be perceived as large hazards.

Some of these unreplicated dimensions appear to result from large grids. Such is the case, for instance, for the "origin" dimensions in the UG sample. Other unreplicated dimension, it was argued, were attributable to the diversity of SA elements. For six of these eight dimensions, however, inter-relations (albeit tenuous at times) with other concepts were pointed out. Thus, it appears that hazard assessment is a very intricate mechanism.

11.4 INTER-RELATEDNESS OF CONCEPTS

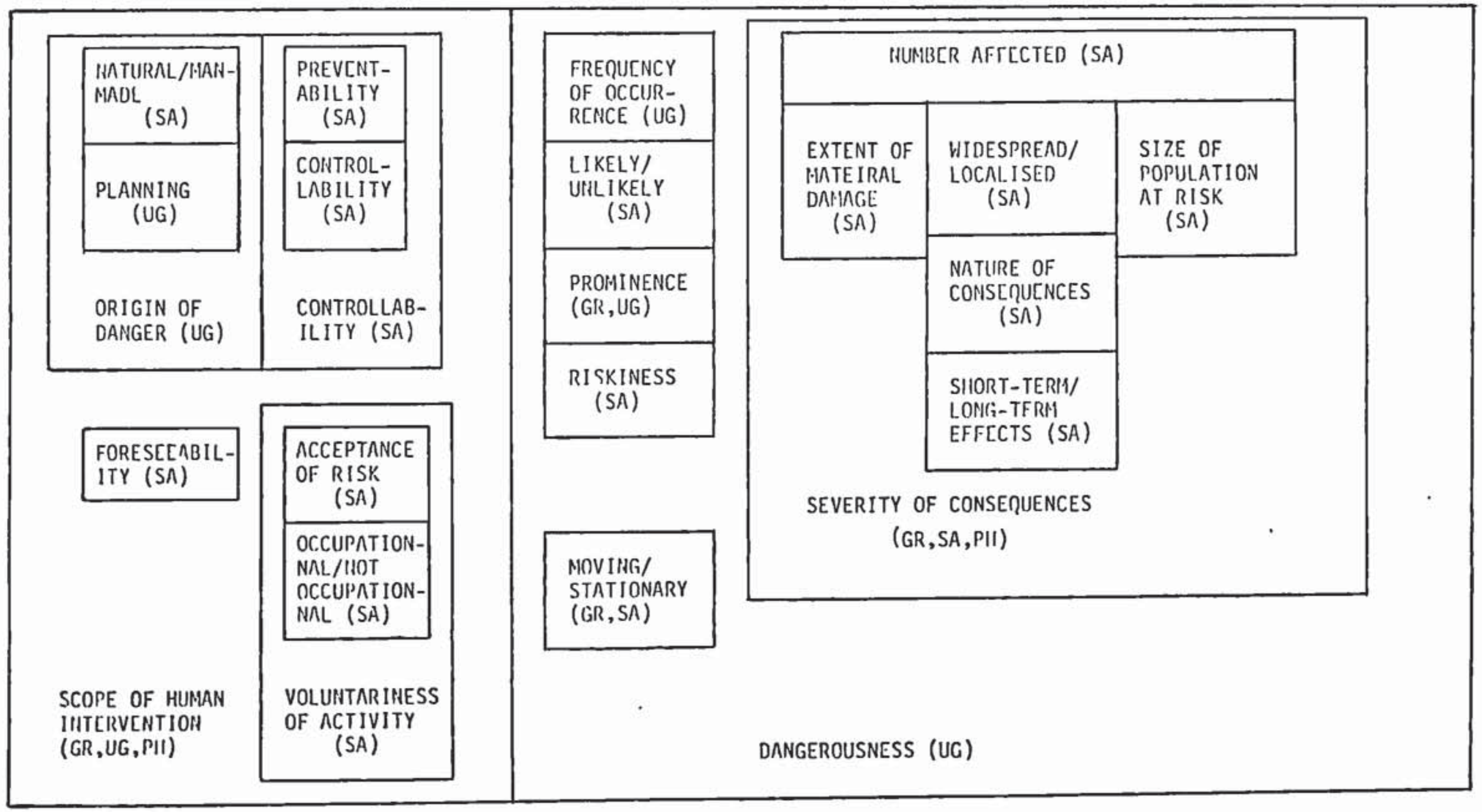
Throughout the review of the main concepts identified in the four samples, many indications of the inter-relatedness of concepts were pointed out. In some cases, it was argued that two dimensions appeared to be strongly correlated. In many other cases, it was mentioned that one dimension reflected a specific aspect of a broader concept.

A systematic description of all the concepts fragmented into their specific facets and of all the plausible correlations between concepts would have to be long and fastidious. Therefore, an attempt has been made to summarize and illustrate these various fragmentations and correlations. The result is Figure 11.1.

A few explanations are necessary in order to understand the figure. Firstly, each box represents a concept as identified in one or more of the samples. The box contains the label used and

FIGURE 11.1

SUMMARY OF FRAGMENTED AND INTER-RELATED CONCEPTS



the sample(s) from which it was derived. All boxes imbedded in larger boxes represent specific facets of these larger concepts. In other words, smaller boxes are meant to represent groups of constructs found among those which form the dimensions represented by larger boxes.

All facets of a concept are assumed to be correlated. There were, however, specific facets of some concepts which showed closer affinities between them than with other facets of the same concepts. These closely associated facets were identified from the fact that they shared at least one construct. In the figure, these contiguous facets are represented by contiguous boxes.

The figure was entitled "summary" for two reasons. Firstly, there were concepts which shared constructs with facets of general dimensions other than their own. For instance, the construct "large/small consequences" was found in both "natural/man-made" and "severity of consequences". These relations across broad concepts are depicted only by the contiguity of the broad concepts. The network of interrelations is somewhat more complex than it appears in Figure 11.1. Secondly, some concepts (e.g. "dread", "familiarity", etc.) are not represented. The notion of "dread", for instance, appears to override the whole of this network, and as such, it is difficult to incorporate it in the figure. Other dimensions, such as "modern/traditional danger" show signs of affinities with both of the most general dimensions in the figure. There is, however, insufficient information about these minor dimensions to incorporate them in their proper place in the figure.

The model depicts two general dimensions. They are "scope of human intervention" and "dangerousness". The first of these dimensions was extracted from three samples. In the last three chapters, it was argued that this dimension encompassed many facets. The model illustrates three major facets ("origin of danger", "controllability" and "voluntariness of activity") and one apparently less important ("foreseeability"). The comparisons within and between samples revealed that these four facets appeared to attract specific sub-sets of constructs from the broader dimension . In turn, sub-sets of constructs from the three larger facets were found in fragmented dimensions. Five of these smaller dimensions ("natural/ man-made", "preventability", "controllability", "acceptance of risk" and"occupational/not occupational") were found in the SA sample and the sixth ("planning") was one of the minor dimensions identified in the UG sample. As can be seen in the model, there were specific indications that some of the facets were inter-related. In some cases, there were significant correlations between the constructs and two smaller dimensions. In other instances, smaller dimensions encompassed some identical constructs.

The second main dimension, labelled "dangerousness", was found in the UG sample. It was argued earlier in this chapter that this UG dimension included constructs reflecting "severity of consequences" (as identified in the other three samples) and the concept of likelihood. Likelihood was expressed in different ways ("frequency of occurrence", "likely/unlikely", "prominence" and "riskiness") in various samples. In all cases, however, there were indications that the facets of the concept of likelihood were correlated with the concept of severity of consequences. The latter concept was fragmented into six specific dimensions in SA grids. Again, these dimensions appeared to be closely inter-related. Finally, "moving/stationary", as an SA dimension and a GR construct, showed affinities with various facets of the concept of "dangerousness".

There are other indications which suggest that the two main dimensions of the model are inter-related. For instance, it was pointed out in Chapter 9 that the first UG dimension had begun to merge into the second dimension. Furthermore, it was mentioned earlier in this section that some minor SA dimensions such as "modern/traditional danger" seemed to be related to both major dimensions. Fischhoff et al's (1978) results suggest that the concept of "modern/traditional danger" belongs to their first factor, labelled "technological risk", which is akin to the "scope of human intervention" dimension identified in this research. However, no conclusive evidence was found on this point. The fact that constructs from both major dimensions were correlated in all samples is further evidence of the inter-relatedness of the main concepts.

The structure of this model suggests a series of hypotheses about the way in which hazard assessment is carried out. Firstly, when only a crude or superficial assessment is required, the first consideration that springs to mind is whether a person can do something about a hazard. If nothing can be done, the next preoccupation is whether a hazard can have severe consequences. The hypothesis that preoccupations emerge in that order is suggested by the order of prominence of the main dimensions in three samples. Order of prominence in the SA sample remains to be determined, as various indications about it were contradictory.

If a more substantial assessment is needed, more specific aspects of intervention and of dangerousness are looked into. As circumstances warrant an increasingly finer assessment, the facets of the two main concepts which enter into consideration also become more specific. Although based on indications found in this research, this description of hazard assessment remains to be tested in research on the behavioural implications of the model.

There were numerous indications that hazard assessment is a complex and intricate process. Nevertheless, Figure 11.1 remains partly hypothetical. For instance, close affinities are depicted between "natural/man-made" from the SA sample and "planning" from the UG sample. This inter-relatedness is inferred from constructs common to the two dimensions. It would be necessary to find both dimensions in one grid in order to test whether the two are indeed correlated. Further research would need to be carried out in order to test the full extent of the model presented in Figure 11.1. In such research, it would be possible to formulate much more precise and articulate hypotheses based on the model than has been possible in the present research. The hazard assessment model postulated at the beginning of this research certainly appears simplistic when compared to the model suggested by the main observations from the results of the four samples.

In the next chapter, the original hypotheses of this research are reviewed in the light of the results discussed in the last five chapters. Subsequently, some suggestions for further research are put forward.

CHAPTER 12 : CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

After discussing and comparing the main results from the four samples, it is necessary to test the accuracy and validity of the original research hypotheses. This review of the hypotheses forms the initial substance of this final chapter.

The use of the repertory grid involved making certain assumptions.. These assumptions are reviewed in order to assess whether the research procedures had an impact upon the main results of this research.

In the light of these discussions, results from the studies described in Chapter 4 are compared again. These studies, while supplying some answers, have also raised many questions. Therefore, the final part of this chapter considers some suggestions for further research and discusses the practical implications arising out of such studies.

12.1 REVIEW OF THE HYPOTHESES

When the results of the first two (GR and UG) samples were discussed, in Chapters 8 and 9, the hypotheses of this research seemed generally to be confirmed. However, after the analysis of the results of the third (SA) and fourth (PH) samples, some of the hypotheses took a serious blow.

12.1.1 First hypothesis

The first hypothesis stated that: "When hazards are assessed as elements of a repertory grid, at least three underlying assessment patterns can be identified by principal component analysis".

Even in the first sample some individual grids yielded only two significant dimensions. In the UG sample, all individual grids, and both mean grids, yielded more than three significant dimensions. In the third sample (SA), three individual grids yielded only one very large dimension. As for the fourth sample, none of its grids generated more than two dimensions; in grids PH01 and PH$\overline{Z}$, there were hardly traces of a second dimension.

In Chapter 11 it was noted that the size of a grid influenced the number of dimensions extracted from it. However, there were also indications that the number of dimensions was a function of the elements being assessed; the more heterogeneous the list of elements was, the more diverse were the assessment patterns and the greater was the variation in the number of dimensions between respondents in a sample. Thus, not only was the prediction of the number of dimensions inaccurate, but the prediction of the nature of the assessment patterns was also incomplete.

12.1.2 Second hypothesis

The second hypothesis was: "The first and most important assessment pattern to be extracted concerns the controllability of hazards".

As already noted, dimensions labelled "controllability" were only extracted from SA grids. Of the five grids which yielded such a dimension, one yielded it in first position, one in second position, two in third place and one in fourth position. Thus, "controllability" as such was not as prominent a concept as expected.

However, in a vast majority of grids in the other samples "controllability" was one of the very important facets of "scope of human intervention". The latter was consistently found in first position.

12.1.3 Third hypothesis

The third hypothesis of this research read: "The second assessment pattern to be extracted concerns the severity of potential consequences of hazards".

In three samples (GR, UG and PH) this hypothesis was confirmed. As pointed out in Chapter 10, however, this hypothesis was not confirmed in the SA sample for two reasons. Firstly, those dimensions labelled "severity of consequences" tended to be more prominent than "controllability" dimensions. Secondly, in a majority of SA grids, the dimensions which were identified dealt with specific aspects of potential consequences rather than with the general concept of "severity of consequences".

12.1.4 Fourth hypothesis

The fourth hypothesis was worded as follows: "The third assessment pattern to be extracted concerns the likelihood of hazards leading to accidents".

This hypothesis was probably the most inaccurate one of this research. As pointed out earlier, no third dimension could be found in any PH grid. No third dimension was common to a majority of SA grids. The same can be said about UG grids, although a third consensus dimension was extracted from both UG mean grids. This third consensus dimension, however, was not identical to its equivalent in the GR mean grids.

In the SA sample, only three dimensions received the label "likely/unlikely". This label was affixed to a first dimension twice and to a fourth dimension once. In the UG grids, constructs dealing with the concept of likelihood were intermingled with others related to "severity of consequences" to form the "dangerousness" dimension.

12.1.5 Fifth hypothesis

The fifth hypothesis stated that: "These assessment patterns are robust and as such they can be identified both by cluster analysis and by principal component analysis".

Since the hypothesized dimensions seldom corresponded to what dimensions were found, it is redundant to discuss their robustness. As for the dimensions which were identified, it has been argued in previous chapters that they were fairly robust. First dimensions were quite systematically robust. Robustness tended to decrease as subsequent dimensions were extracted. In the three samples for which consensus grids were computed, individual dimensions which had a namesake in mean grids were more robust than those dimensions which were found only in a few individual grids. In the SA sample, the less robust dimensions tended to be those which encompassed constructs which elicited many ratings of zero.

12.1.6 Sixth hypothesis

The sixth hypothesis was that: "Most individual grids having identical elements yield identical hazard assessment patterns".

The main individual dimensions in the GR, UG and PH sample were comparable between respondents. However, in the UG sample, there were quite a few minor dimensions which varied from one grid to another. In the SA sample, this hypothesis proved to be totally erroneous. For instance, no two grids yielded identical series of dimensions. A few pairs of grids yielded similar dimensions but in different orders. As pointed out in Chapter 10, three SA grids yielded only one dimension each. These three dimensions shared constructs about repulsiveness of hazards and about dread; but the fact that they were labelled differently implies that there were connotations (such as "panic/near-calm" in one grid) which were different from one dimension to another.

As pointed out in Chapter 11, the nature of the elements did influence hazard assessment patterns. This influence, however, was not of the type postulated in the sixth hypothesis. In fact, it was found that the heterogeneity within a list of elements led to heterogeneous assessment patterns among the respondents of a sample. It appears that a wide range of elements generated a great diversity of constructs. Since a construction system, according to Kelly (1955), includes a limited number of constructs, the constructs sampled from a wide range by various respondents formed different sets. In turn, different sets of constructs yielded different dimensions. Thus, whilst this hypothesis was partly confirmed in three samples, it was found to be erroneous in the SA sample.

12.1.7 Seventh hypothesis

The seventh hypothesis formulated in Chapter 5 dealt with consensus dimensions. It read: "Hazard assessment patterns extracted from consensus grids are similar across samples".

It was argued in the previous chapter that the first two consensus dimensions in the GR, UG and PH samples were very similar. It was also pointed out, however, that the second UG consensus dimension was more general in nature than its equivalent in the other two samples; "severity of consequences" was only one of the two major facets of the concept of "dangerousness".

There were, however, dissimilarities in consensus dimensions even between these three samples. For instance, there were two consensus dimensions in PH mean grids, four in GR mean grids and five in UG mean grids. *Furthermore, the additional GR and UG dimensions* had very little in common.

No concensus grid could be computed for the SA sample. If such a grid had been computed, in view of the diversity of the results in that sample, it seems unlikely that the consensus dimensions would have been identical to those of any other sample. Nevertheless, because of apparent correlations between specific dimensions, it seems plausible that consensus dimensions reflecting "scope of human intervention" and "dangerousness" would have emerged from SA consensus grids. It was pointed out in Chapter 10 that all dimensions reflecting specific aspects of "scope of human intervention" shared common constructs. All dimensions concerned with specific aspects of "severity of consequences" also showed some common contents. Together these two sets of dimensions were the most prominent in the large majority of individual grids. Hence, it would seem likely that corresponding consensus dimensions could have been extracted if consensus grids could have been computed.

To a certain extent, it can be said that the notions of "scope of human intervention" and "dangerousness", along with their specific facets, formed the core of hazard assessment throughout the four samples. That is why the model proposed in Chapter 11 rests essentially on two general dimensions. Strictly speaking, however, the consensus dimensions were not totally identical across samples.

12.1.8 Eighth hypothesis

The eighth and final hypothesis of this research postulated individual differences between grids. It was formulated as follows: "Whilst emerging dimensions are identical across the grids of a sample, hazards can be construed differently by different respondents and hence ratings of elements on constructs are not unanimous".

The discussion of the previous hypotheses has highlighted the fact that the first part of this eighth hypothesis was not accurate. In other words, it cannot be said that "emerging dimensions are identical across the grids of a sample". Thus, individual differences between grids were more fundamental than mere rating differences. In fact, different dimensions between grids suggest different construing and not only different ratings.

Seventeen of the 54 cells in the PH grid received unanimous ratings. Comparisons of ratings could not be made systematically in the SA sample. However, there are indications that on some constructs which were common to a few grids hazards were not rated unanimously. As for the grids of the other two samples, despite some high rating frequencies, no element received an identical rating on a construct in all grids of a sample. Thus, it seems plausible that unanimous ratings in the PH sample are attributable to the fact that there were only three PH respondents. Nevertheless, the possibility that high sample homogeneity led to unanimous ratings cannot be excluded.

Evidence for testing this hypothesis was not thoroughly examined in the previous chapters. Differences found between grids were more important and deserved more attention than specific ratings. Different analyses would be needed in order to study the behaviour of elements across grids. Further research should be undertaken on this point.

12.2 REVIEW OF GRID ASSUMPTIONS

There were good reasons, mentioned in Chapter 5, for choosing the repertory grid as the measurement instrument for this thesis.

In using the grid, however, certain assumptions had to be made. These assumptions were identified by Kelly (1955). They were also discussed in Chapter 5.

It was pointed out then that some assumptions could not be tested or controlled. For instance, it was not within the scope of this thesis to test the assumption that the construction system identified through a grid serves as a basis for a respondent's behaviour. Although the aim of this research was not to test the other assumptions either, some strategies were divised in order to gain some insight into the validity of these assumptions in this research.

The first assumption is that elicited constructs are permeable, i.e. that they can be used for assessing elements other than those from which they were elicited. Constructs such as "controllable/not controllable" and "likely to kill/unlikely to kill" were used in all samples. These constructs can therefore be considered as permeable.

In three samples (GR, UG and PH) all respondents used a five-point rating scale. In the SA sample, respondents were also given the possibility of assigning a rating of zero, thus signifying that a construct was irrelevant to the assessment of an element. It was noticed that the vast majority of zero ratings were found on constructs used only in the SA sample (such as "painful/painless", "difficult rescue/more immediate rescue"). Constructs which do not seem to be truely bi-polar (e.g. "going-to-bed type situations/entertainment") tended to generate seven or more ratings of zero out of a possibility of twenty-one ratings.

The use of zeros was discussed in Chapter 10. It was argued then that zeros introduced little or no bias in the extraction of dimensions. Thus, although some impermeable constructs were used, they tended to be minor constructs which had little effect on the main results of this research. The more permeable constructs tended to be the main themes of the dimensions discussed in Chapter 11.

Another assumption underlying a repertory grid was that the elements of a grid were representative of all the elements within their category. In this research, three different types of hazards (home hazards, occupational hazards and pot-holing hazards) led to the identification of largely comparable dimensions. It was argued in the previous chapter that SA elements led to fragmented dimensions. Thus, GR, UG and PH elements may not be representative of major disasters. Beyond fragmentation, however, there are indications that SA dimensions were akin to those from other samples.

The use of a repertory grid also rests on the assumption that constructs are not "concocted on the spot" (Kelly, 1955). This assumption could not be verified in this research. However, the fact that the most important concepts re-emerged in all samples makes it unlikely that the corresponding constructs should be artefacts of the grid method. It was argued in Chapter 5 that repertory grids, like other measurement instruments, rest on the assumption that all the important facets of a person's construction system are elicited. There are a number of indications that this assumption was valid. Firstly, the vast majority of dimensions were robust. Therefore, the important dimensions are not likely to have been statistical artefacts. Secondly, in all grids the significant dimensions explained approximately 80% of common variance. In contrast, Golant and Burton's (1968) dimensions explained less than 50% of common variance. It was argued in Chapter 4 that a low level of explained variance indicated the absence of important concepts. Finally, group discussions were carried out in two samples (UG and PH). These discussions did not lead to the identification of important dimensions not found in the other two samples.

To a certain extent, it was also possible to verify the assumption that elicited constructs were "functionally communicable" (Kelly, 1955) i.e. that the meaning of a person's constructs can be understood by someone else. In all samples, key concepts had the same connotations. In other words, pairs of constructs found in more than one sample tended to correlate in the same way in all samples. Furthermore, constructs elicited during group discussions required very little clarification. Thus, these constructs seemed to be understood by all respondents.

The arguments discussed above do not constitute a thorough validation of the assumptions. Nevertheless, they remove some of the uncertainty about the assumptions. Therefore, proportionately greater confidence can be lent to the main results of this research.

12.3 CONCLUSIONS

Probably the most plausible conclusion which must be drawn from this brief review of the hypotheses is that the results of this research were far more complex than originally foreseen. Thus, in retrospect, the hypotheses appear to over-simplify the results.

These hypotheses, it will be remembered, were formulated following a critical review of other studies of hazard assessment. Thus, if the hypotheses were incorrect, it was either because the critical review was not carried out systematically or because the evidence available from previous studies was incomplete.

The possibility that the analysis of previous studies was not systematic cannot be rejected completely. There are, however, indications that the results of the studies reviewed in Chapter 4 were not completely reliable. For instance, when those results were discussed, a number of methodological shortcomings were pointed out in each of the studies reviewed. Furthermore, in the light of the results of this research, a number of additional shortcomings can be pointed out which were not obvious at first.

For instance, the studies described in Chapter 4 were each based upon the results from only one sample. In this research, if the results from only the GR sample had been used for example, then the conclusions of this research would have been a lot less elaborate than they are.

By restricting their data collection to only one sample, the researchers whose studies are reviewed in Chapter 4 each used only one list of elements. Since this list varied from one study to another, and probably because the degree of homogeneity among elements varied from one list to another, some aspects of the results differed from one study to another. The results of this research suggest that hazard assessment involves complex mechanisms and that different lists of elements shed light on different facets of these mechanisms.

Two of the studies quoted in Chapter 4 (Golant and Burton, 1969; Champion, 1977) relied mainly on the use of a consensus grid. The results of this research suggest that a consensus grid is a fairly accurate method of summarizing the main trends within a sample. However, some important phenomena would have been overlooked by sole reliance upon a consensus grid in this research. In the UG sample, for instance, consensus dimensions alone would have masked some meaningful differences between individual grids. If a consensus SA grid has been computed and used as the only source of results, even more important differences between individual grids would have been overlooked and the hazard assessment structure discussed in Chapter 11 would not have been as complex as it is. It was pointed out in Chapter 4 that Golant and Burton's dimensions were somewhat difficult to interpret. In the same chapter, it was argued that confusion stemmed from the researchers' choice of scales. Six of their elements were also found in SA grids. In view of the diversity of SA dimensions, it may be that Golant and Burton's consensus dimensions were but a sketchy summary of phenomena which individual questionnaires would have revealed to be very complex.

The data used by Green and Brown (1976a, 1976b) and by Champion (1977) were thoroughly re-analyzed in this research (Chapters 10 and 8 respectively). These results, along with those from two other samples, revealed that the hypothesized dimensions could be identified, but that these dimensions were part of a more complex mechanism.

Thus, the hypothesized "controllability" dimension, similar to that revealed by Golant and Burton (1969), appears to be but one of the facets of a more general dimension. The latter was labelled "scope of human intervention"; its contents were comparable to Fischhoff et al's (1976) "technological" dimension.

The second ("severity of consequences") and third ("likelihood of occurrence") hypothesized dimensions appear to belong within the second general dimension, labelled "dangerousness". The dimension occasionally found in this research and labelled "severity of consequences" was very similar to Fischhoff et al's "dread" dimension; for instance, in both dimensions the main construct was "likelihood of death". Golant and Burton's "magnitude" dimension encompassed many scales found as constructs in the "widespread/ localized" dimension in the SA sample. There were also similarities between the "likelihood of occurrence" dimension found in the SA sample and the "expectancy" dimension suggested by Golant and Burton. Thus, "dangerousness" seems to encompass dimensions found in other studies.

The model derived from the results of this research and presented in Chapter 11 suggests that the two general dimensions ("scope of human intervention" and "dangerousness") are inter-related. This is substantiated in Fischhoff et al's (1976) research where the two dimensions have two significant constructs in common.

There are indications, therefore, that the model proposed in Chapter 11 does encompass results from other studies as well as the results of this research. The model can probably serve as a very useful basis for future research. However, there is nothing in the model which can be compared to Golant and Burton's main dimension ("stability"). It may be that Golant and Burton's dimension is a statistical artefact. It may also be that, since "pleasant/unpleasant" is an important facet of "stability", this dimension is akin to the "dread" dimension which appears to take precedence in three SA grids. In any case, further research is needed in order to define some aspects of the model more precisely.

12.4 SUGGESTIONS FOR FURTHER RESEARCH

There are many potential areas for future research. For instance, as pointed out earlier, some aspects of the model remain to be tested and other aspects need to be formulated in more specific terms. There are also behavioral implications to hazard assessment; it might prove very useful to investigate these implications. These potential areas for future research are discussed separately in this section.

12.4.1 Research on the model

It was mentionned in Chapter 11 that many of the correlations postulated by the model were hypothetical; these correlations could not be tested since the dimensions involved were not found in the same grids. It might be possible to design experiments in order to test these correlations. For instance, an experiment could be designed in which a grid would be constituted so as to include the necessary constructs and be submitted to respondents. The same respondents would also have to rate individual grids in order to test the level of concordance between the pre-arranged grid and the individual grids.

Since the model postulates different levels of generality (or conversely, of specificity), it might be interesting to check whether these different levels correspond to different strengths of correlations. Principal component analysis performed on component scores would be one way of investigating the various levels of specificity or of differentiated thinking.

Other research strategies could also prove to be very worthwhile. For instance, individual and consensus grids using Golant and Burton's elements might shed some light on these authors' results and on the model proposed in this research. Another strategy could consist of double rating; in other words, respondents within a sample could rate a list of elements on their own constructs and later rate the same elements on the full set of constructs elicited within the sample. Such a procedure could highlight the effects of imposing a set of constructs upon respondents; the differences in assessment patterns for the two rating procedures could also help to define the model more accurately.

Throughout this research, a number of individual characteristics of respondents which were different between samples have been pointed out. Age, sex, education (and cognitive complexity), origin (and knowledge of English), marital status (and number of children), knowledge and experience in health and safety matters were all respondent characteristics on which some samples differed. There were also differences between the grid techniques (e.g. order of presentation, construct elicitation technique, etc.) used in the four samples.

This research was not designed to rigorously control these variables. Different experimental designs would be needed in order to perform a systematic test of the potential influence of these variables. Nevertheless, some comparisons were possible. The potential influence of these individual and technical variables on some major aspects of the results was tested for and always found to be minimal. Therefore, it would appear that studies on the nature of hazards rather than individual or methodological variables should be turned to in order to pursue the development of a theory of hazard assessment.

The results of this research suggest that the generality or specificity of dimension labels may be mainly situational, i.e. influenced by the nature of the hazards being assessed. For instance, there seem to be hazards for which the notion of control is somewhat irrelevant. For the major hazards in the SA sample, the severity of potential consequences is analyzed in more detail. This may suggest that within the general framework identified in this research, the specific mechanisms which are activated depend upon the nature of the hazards being considered.

The data at hand can be submitted to a few more analyses, but these analyses are likely to be incomplete. Specially designed experiments should be carried out to investigate the reasons why dimensions are reversed in some individual grids. For instance, one experiment should focus specifically on major hazards in order to avoid the type of fragmentation found in the SA sample. Another experiment should ask respondents to assess various types of hazards in successive grids, in order to make sure that it is different facets of the same mechanism which are activated by different types of hazards.

12.4.2 Research on the implications of hazard assessment

The various possibilities discussed above indicate that much research can be carried out on the mechanisms of hazard assessment. But it is also possible to look into the behavioral connotations of hazard assessment. For instance, the grids discussed in this research were rated (and the hazards assessed) in the relative safety of either the schoolroom or the home. When hazards are encountered in reality, however, it is not certain whether ratings, assessment patterns and priorities remain unchanged.

It might prove interesting and worthwhile to find out whether an emotional reaction (such as that which is thought to have occurred in three SA grids) is indicative of a tendency to panic in the face of danger. The results of this research suggest that there can be individual differences in hazard assessment patterns. The reasons for these differences are not obvious. Is hazard assessment influenced by accident experience? Can it be shown that those persons who have a tendency to take positive action in the face of dangers have different hazard assessment characteristics? Are there assessment patterns which characterize efficient factory inspectors or safety officers? What is the effect of training programs aimed at dealing with specific hazards? Do motivation factors (e.g. first aid training) alter the way in which hazards are assessed?

At present, it cannot be stated whether assessment dimensions develop into a fixed mechanism or whether these dimensions can be changed. There may be a parallel worth exploring between attitude formation and change on the one hand, and hazard assessment mechanisms on the other. It might be worthwhile finding out whether training or tuition in health and safety bring about a change in hazard assessment. For instance, the hypothesis can be made that, when students complete a university course in health and safety, their hazard assessment patterns are different, more articulate than when they started the course.

12.5 SUMMARY

In addition to those already mentioned, many more questions could be formulated about the nature, development and implications of hazard assessment mechanisms. But it seems very likely that many of the answers can be found only if there is a theoretical framework within which hypotheses can be clearly and accurately stated. The results of this research led to the elaboration of a model which constitutes a useful basis upon which to structure further research.

The results of previous studies were reviewed in Chapter 4. It was pointed out then that despite some similarities these results were not unanimous. It was argued that differences in results could be attributable to differences either in research methods, in respondent characteristics, or in the specific aspects of hazards being studied.

The findings of this research revealed that the original hypotheses based on previous studies were over-simplified. A complex model, presented in Chapter 11, was elaborated to depict the important results from the four samples. Since this research was not designed to validate such a complex model, some of its facets remain to be tested. Nevertheless, many indications suggest that the model is a relatively accurate representation of the hazard assessment structure.

Firstly, the model is an integration of results obtained from samples in which different research strategies were used. A number of methodological and technical variables were controlled. These variables were not found to significantly influence the main results. In this respect, this research removes some of the ambiguities arising from previous studies.

Secondly, the dimensions in the model were extracted from grids rated by different types of respondents. Some individual respondent characteristics were different either within or between samples. Yet, the identified dimensions could not be found to be affected by these differences in respondent characteristics. Therefore, lack of unanimity among the results from previous studies could not be explained by differences between their respondents.

Thirdly, the results from this research as presented in the model were derived from a wide range of hazards being assessed on a large number of characteristics. Beyond the diversity of constructs and elements, dimensions and their facets were similar across samples. Thus, although different parts of the model may be more relevant than others in different situations, it is very likely that the model is broad enough in scope to account for the assessment of diverse hazards.

Probably because it was elaborated on such a broad basis, the model encompasses (and is substantiated by) the results from previous studies. Thus, the most plausible explanation for the differences between these previous results appears to be that previous studies were concerned with different hazards and different hazard characteristics. Therefore, there are substantial reasons for believing that the model which depicts the main results of this thesis can be a cornerstone for a theory of hazard assessment as well as a firm basis for future research.

APPENDIX 1

LIST OF ELEMENTS USED IN CHAMPION'S (1977) REPERTORY GRIDS

1. Main valves left open on oxy/acetylene welder
2. Donkey jacket draped across convector heater
3. Emergency doors blocked-out for film show
4. Pipes across walkway
5. Badly fitting grid sticking up above floor level
6. Electric cable lying on floor
7. Tablet of soap on wet floor
8. Gas cylinder left free-standing on ramp
9. Scaffolding stacked in dangerous manner
10. Forklift truck being driven without warning light
11. Guard missing off machinery
12. Beer on floor makes it slippery
13. Operator leaning across conveyor to operate machinery
14. Steam gushing out of tanker where people walk
15. Cut-off electric eye not working
16. Oil spillage
17. Kegs thrown off conveyor by faulty reject arm
18. Keg lift faulty so operator has to brake hard
19. Bad ladder footing

APPENDIX 2

LIST OF CONSTRUCTS USED IN CHAMPION'S (1977) REPERTORY GRID

1.	Very often encountered in my job	- Never encountered in my job
2.	Necessary result of the process	- Danger not a necessary result of the process
3.	Temporary danger	- Permanent danger
4.	Moving danger	- Stationary danger
5.	Easily spotted danger	- Very difficult to identify danger
6.	Danger is immediately present	- Danger is dependent upon other things
7.	Very likely to cause an accident	- Very unlikely to cause an accident
8.	Easy to avoid consequences of danger	- Impossible to avoid consequences of danger
9.	Only a trivial injury	- Permanent disability
10.	Very likely to kill	- Very unlikely to kill
11.	Only one person at risk	- Every person in the plant at risk
12.	Preventable	- Unpreventable
13.	Takes a specialist to put it right	- Anyone can put it right
14.	Danger arises from bad design feature	- Danger has nothing to do with design
15.	Management's fault	- Nothing to do with management
16.	Operator's fault	- Nothing to do with operator
17.	Due to inadequate training	- Danger has nothing to do with training

APPENDIX 3

LIST OF HAZARDS USED AS ELEMENTS FOR THE GRID

RATED BY UG RESPONDENTS

1. Too many things going on at once.
2. Ball in the way.
3. Rucked up mat.
4. Soap suds on the floor.
5. Bucket in the way.
6. Knife sticking out of the drawer.
7. Tin left on the floor.
8. Jagged edge on tin lid.
9. Flip flop shoes.
10. Matches on cooker.
11. Kettle lead in water.
12. Faulty wire on mixer.
13. Confusion over on and off position of switches.
14. Fat boiling over.
15. Mixer overhanging the working surface.
16. Cooker switches hard to reach.
17. Two (2) cooker rings on.
18. Kettle (electric) left on cooker ring.
19. The fact that the drawer is open.
20. *Lady not looking at what she is doing.*
21. The way the lady is holding the vegetable for cutting.
22. Improperly stacked bottles in cupboard.
23. Door of cupboard is open.
24. Position of the bin.
25. The fact that the bin is overfilled.
26. Broken bottle sticking out of the bin.
27. The man's vision being obscured by the box.
28. Box with fragile contents, improperly sealed.
29. The way the man is holding the box.
30. Switch on mixer left "on".
31. Untidy work surface (loose vegetables, etc.)
32. Lady not wearing an apron.
33. Tin opener left lying about.
34. Edge of botton cupboard.
35. Edge of top cupboard.
36. Stacking of cups and saucers.
37. Position of cooker as regards opening of oven door.
38. Position of the chip pan handle.

APPENDIX 4

LIST OF CONSTRUCTS AND THEIR SCALE IN THE UG GRID

CONSTRUCT NUMBER	DESCRIPTION	SCALE
1	This is the consequence of another hazard	Strongly agree / Strongly disagree
2	This is a hazard	Strongly agree / Strongly disagree
3	This is due to inadequate maintenance	Strongly agree / Strongly disagree
4	This hazard is only temporary	Strongly agree / Strongly disagree
5	This is the result of untidiness	Strongly agree / Strongly disagree
6	In order to call this a hazard, one has to make some assumptions	Strongly agree / Strongly disagree
7	This is caused by a lack of knowledge	Strongly agree / Strongly disagree
8	This is sufficient on its own to cause an injury	Strongly agree / Strongly disagree
9	This is the result of hurrying	Strongly agree / Strongly disagree
10	It takes a specialist to put this right	Strongly agree / Strongly disagree
11	This arises from something that is in use	Strongly agree / Strongly disagree
12	This is the result of poor planning	Strongly agree / Strongly disagree
13	This is a multiple hazard	Strongly agree / Strongly disagree
14	This arises from something not having been done	Strongly agree / Strongly disagree
15	This is due to bad design	Strongly agree / Strongly disagree
16	This is the result of someone's action(s)	Strongly agree / Strongly disagree

APPENDIX 4 (continued...)

No.	DESCRIPTION	SCALE
17	How hard/easy is it to see this?	Very easy / Very hard
18	How easy/difficult is it to remove this?	Very difficult / Very easy
19	How easy/difficult is it to avoid this?	Very difficult / Very easy
20	How easy/difficult is it to put this right?	Very difficult / Very easy
21	How serious/trivial an injury could this cause?	Very trivial / Very serious
22	How likely/unlikely is this to cause an accident?	Very unlikely / Very likely
23	How direct/indirect a cause of injury can this be?	Very indirect / Very direct
24	How large/small a hazard is this?	Very small / Very large
25	How dangerous is this?	Very dangerous / Not dangerous at all
26	One comes across this...	Very frequently / Very infrequently
27	How likely/unlikely is this to cause death?	Very likely / Very unlikely

APPENDIX 5

LIST OF ELEMENTS USED BY GREEN AND BROWN (1976 a) IN EXPERIMENT E1

1. Hotel fire
2. Coal mining accident
3. Air pollution
4. Fire in a discotheque
5. Illness
6. Train crash
7. Accidental release of nuclear radiation
8. Home fire
9. Plane crash
10. Earthquake
11. Car crash
12. Rock-climbing accident
13. Accident on a building site
14. Being struck by lightning
15. Accident in a chemical plant
16. Skiing accident
17. Factory fire
18. Accident in the home
19. Food poisoning
20. Motorcycle crash
21. Being knocked down while crossing the road

APPENDIX 6

LIST OF HAZARDS USED AS ELEMENTS BY PH RESPONDENTS

1. Failure to take space lighting on long trip.
2. Insufficient care taken when climbing dislogded rocks.
3. Dislodged miners' deads in a mine.
4. Digging in an unstable choke without shoring.
5. Bad air in airbell between sumps.
6. Inexperience of party especially leader.
7. Rope abrasion on SRT rope.
8. Loose rocks in boulder choke.
9. Poorly supervised novices.

APPENDIX 7

LIST OF CONSTRUCTS USED BY PH RESPONDENTS

1.	Controllable	-	Uncontrollable
2.	Lack of thought	-	Bad luck
3.	Due to my party	-	Due to other party
4.	Technical factor	-	Physical factor
5.	Poor planning	-	Poor moving
6.	Likely to kill	-	Unlikely to kill

REFERENCES

ABEYTUNGA, P.K. 'Construction safety: the role and training needs of the supervisor' PhD dissertation, Department of Safety and Hygiene, University of Aston in Birmingham, 1979.

AJZEN, I. and FISHBEIN, M. 'Use and misuse of Bayes' theorem in causal attribution' Psychological Bulletin, 85(2): 244-246, 1978.

ALDRIDGE, J.F.L. 'Safety and Motivation' Safety Surveyor, September, 19-25, 1976.

ANDERSSON, R., JOHANSSON, B., LINDEN, K., SVANSTROM, J.and SVANSTROM, L. 'Development of a Model for Research on Occupational Accidents' Journal of Occupational Accidents, 1: 341-352, 1978.

ATHERLEY, G,R.C., HALE, A.R. and McKENNA, S.P. 'The prevention of accidents at work: the role of first aid training' Royal Society of Health Journal, 93(4): 192-202, 1973.

BANNISTER, D. and MAIR, J.M.M. 'The evaluation of personal constructs' London: Academic Press, 1968.

BECKER, M.H. and MAIMAN, L.A. 'Sociobehavioral determinants of compliance with health and medical care recommendations' Medical Care, 13: 10-24, 1975.

BECKER, M.H., MAIMAN, L.A., KIRSCHT, J.P., HAEFNER, D.P. and CRACHMAN, R.H. 'The Health Belief Model and Prediction of Dietary Compliance: a Field Experiment' Journal of Health and Social Behavior, 18: 348-366, 1977.

BIERI, J. 'Cognitive Complexity - Simplicity and Predictive Behaviour' Journal of Abnormal and Social Psychology, 51: 263-268, 1955.

BURT, C. and BANKS, C. 'A Factor Analysis of Body Measurements for British Adult Males' Annals of Eugenics, 13: 238-256, 1947.

BURTON, I., KATES, R.W. and SNEAD, R.E. 'The human Ecology of Coastal Flood Hazard in Megalopolis' Research Paper No 115, Department of Geography, University of Chicago, 1969.

CATTELL, R.B. 'Factor analysis' New York: Harper, 1952.

CHAMPION, A. 'Perception of Danger' MSc dissertation, Department of Safety and Hygiene, University of Aston in Birmingham, 1977.

CHAMPION, A. Personal communication. 1978.

CHILD, D. 'The Essentials of Factor Analysis' London: Holt, Rinehart and Winston (4th Edition), 1976.

COHEN, J. 'Behaviour in Uncertainty' London: George Allen & Unwin, 1964.

COHEN, J. 'Uncertainty and Risk-taking in Crime' Bulletin of the British Psychological Society, 23: 293-296, 1970.

COHEN, J. 'The Nature of Gambling' Scientia, 105(699-700): 1-25, 1970.

COHEN, J. and CHESNICK, E.I. 'The Doctrine of Psychological Chances' British Journal of Psychology, 61(3): 323-334, 1970.

COHEN, J. and CHRISTENSEN, I. 'Information and Choice' Edinburgh: Oliver and Boyd, 1970.

COHEN, J. BOYLE, L.E. and SHUBSACHS, A.P.W. 'Rouge et noir: Influence of Previous Play on Choice of Binary Event Outcome and Size of Stake in a Gambling Situation' Acta Psychologica, 31: 340-352, 1969.

COHEN, J., DEARNALEY, E.J. and HANSEL, C.E.M. 'The Risk Taken in Crossing the Road' Operational Research Quarterly, 6(3): 120, 1955.

COHEN, J. DEARNALEY, E.J. and HANSEL, C.E.M. 'Risk and Hazard' Operational Research Quarterly, 7(3): 67, 1956.

COHEN, J. DEARNALEY, E.J. and HANSEL, C.E.M. 'The Risk Taken in Driving Under the Influence of Alcohol' British Medical Journal, 1: 1438, 1958.

DUNN, J. 'Safety psychology: A Review of the Literature' A.P. Report No 35, Department of Applied Psychology, University of Aston in Birmingham, 1971.

EDWARDS, A.L. 'Techniques of Attitude Scale Construction' New York: Appleton-Century-Crofts, 1957.

EDWARDS, W. and SLOVIC, P. 'Seeking Information to Reduce the Risk of Decisions' American Journal of Psychology, 78(2): 188-197, 1965.

EDWARDS, W., SLOVIC, P. and LICHTENSTEIN, S. 'Boredom-induced Changes in Preferences among Bets' American Journal of Psychology, 78(2): 208-217, 1965.

FISCHHOFF, B., SLOVIC, P., LICHTENSTEIN, S., READ, S. and COMBS, B. 'How Safe is Safe Enough? A Psychometric Study of Attitudes towards Technological Risks and Benefits' Policy Sciences, 9: 127-152, 1978.

FRANSELLA, F. and BANNISTER, D. 'A manual for Repertory Grid Technique' London: Academic Press, 1977.

GLENDON, A.I. 'Safety Economics' Occupational Health, 27(11): 488-490, 1975.

GOELLER, B.F. 'Modelling the traffic-safety system' Accident Analysis and Prevention, 1: 167-204, 1969.

GOLANT, S. and BURTON, I. 'The Meaning of a Hazard-application of the Semantic Differential' Natural Hazard Research Working Paper No 7, Department of Geography, University of Chicago, 1969.

GRAY, J.A. 'The Psychology of Fear and Stress' London: Weidenfeld and Nicolson (World University Library), 1971.

GREEN, C.H. 'Measures of Safety' Urbana: Center for Advanced Studies, University of Illinois, 1974.

GREEN, C.H. 'Questionnaire M2 - Results of Pretest' Dundee: School of Architecture, University of Dundee, 1975.

GREEN, C.H. 'To Determine a Measure of Safety and Aversion as a Function of Safety: a Proposal' Dundee: School of Architecture, University of Dundee, 1975.

GREEN, C.H. 'Measuring Life Safety' Paper to EDRA 7 Conference, Vancouver, 1976.

GREEN, C.H. Personal communication. 1978.

GREEN, C.H. and BROWN, R.A. 'The Perception of, and Attitudes towards Risk: Interim Report' Dundee: School of Architecture, University of Dundee, 1976.

GREEN, C.H. and BROWN, R.A. 'The Perception of, and Attitudes towards Risk: Preliminary Report: E2 Measure of Safety' Dundee: School of Architecture, University of Dundee, 1976.

GREEN, C.H. and BROWN, R.A. 'The Perception of, and Attitudes towards Risk: E3 Stability of Perception under Time and Data' Dundee: School of Architecture, Univer[illegible] of Dundee, 1977.

GREEN, C.H. and BROWN, R.A. 'The Perception of, and Attitudes towards Risk: Final Report: E2 Measure of Safety' Dundee: School of Architecture, University of Dundee, 1977.

GREEN, C.H. and BROWN, R.A. 'The Perception of, and Attitudes towards, Risk: Final Report: E4 Initial Experiments on Determining Satisfaction with Levels of Safety' Dundee: School of Architecture University of Dundee, 1977.

GREEN, C.H. and BROWN, R.A. 'The acceptability of Risk: First Report: Pretesting' Dundee: School of Architecture, University of Dundee, 1977.

GUILFORD, J.P. and FRUCHTER, B. 'Fundamental Statistics in Psychology and Education' New York: McGraw-Hill (5th Edition), 1973.

HALE, A.R. 'Appreciation of Risk at Work' Bulletin of the National Institute of Industrial Psychology, Spring: 27, 1969.

HALE, A.R. and HALE, M. 'Accidents in Perspective' Occupational Psychology, 44: 115-121, 1970.

HALE, A.R. and HALE, M. 'A Review of the Industrial Accident Research Literature' London: HMSO, 1972.

HALE, A.R. and PERUSSE, M. 'Attitudes to Safety: Facts and Assumptions' Social Sciences Research Council Safety at Work Conference. Oxford: Centre for Socio-Legal Studies, Wolfson College, 1977.

HALE, A.R. and PERUSSE, M. 'Perception of danger: a Prerequisite to Safe Decisions' Institute of Chemical Engineers Symposium Series No 53, Birmingham, 1978.

HARMAN, H.H. 'Modern Factor Analysis' Chicago: University of Chicago Press, 1967.

JARRY, J.J., CAZAMIAN, P., CHICH, Y., DEVEZE, G. and FAURE, G. 'Ergonomics and Safety' Archives des Maladies Professionnelles, 23(6): 363, 1962.

KAISER, H.F. 'Computer Program for Varimax Rotation in Factor Analysis' Educational and Psychological Measurement, 19: 413-420, 1959.

KATES, R.W. 'The Perception of Storm Hazard on the Shore of Megalopolis' Research Paper No 109, Department of Geography, University of Chicago, 1967.

KELLY, G.A. 'The Psychology of Personal Constructs. Volume 1: a Theory of Personality' New York: W.W. Norton, 1955.

KIMBLE, G.A. 'Hilgard and Marquis'Conditioning and Learning' New York: Appleton-Century-Crofts, 1968.

KLAUSNER, S.Z. (ed.)'Why Man Takes Chances: Studies in Stress-Seeking' Garden City: Doubleday, 1968.

LAWRENCE, A.C. 'Human Error as a Cause of Accidents in Gold Mining' Journal of Safety Res ch, 6(2): 78-88, 1974.

LESAGE, P.B. 'Measuring Leadership Attitudes: a Construct Validation Study' PhD dissertation, University of Michigan, 1973.

LICHTENSTEIN, S. 'Bases for Preferences among Three-outcome Bets' Journal of Experimental Psychology, 69(2): 162-169, 1965.

LIKERT, R. 'A technique for the Measurement of Attitudes' New York: Appleton-Century-Crofts, 1957.

MASLOW, A.H. 'Motivation and Personality' New York: Harper & Row, 1954.

McKENNA, S.P. 'The Effects of First Aid Training on Safety: a Field Study of Approaches and Methods' PhD dissertation, Department of Safety and Hygiene, University of Aston in Birmingham, 1978.

MELINEK, S.J., WOOLLEY, S.K.D. and BALDWIN, R. 'Analysis of a Questionnaire on Attitudes' Fire Research Note No 962, Borehamwood: Fire Research Station, 1973.

MERRIMAN, R.J. 'The Role of Audiometry in the Prevention of Occupational Deafness' PhD dissertation, Department of Safety and Hygiene, University of Aston in Birmingham, 1978.

MERZ, F. 'Individual Predisposition to Risk-taking' in 'Human Factors and Safety in Mines and Steel Works'. Studies in Physiology and Psychology at Work No 2. Luxembourg: Communauté Européenne de charbon et d'acier, 1967.

MILLER, G. and AGNEW, N. 'First Aid Training and Accidents' Occupational Psychology, 47: 209-218, 1973.

OSGOOD, C.E., SUCI, G.J. and TANNENBAUM, P.H. 'The Measurement of Meaning' Chicago: University of Chicago Press, 1957.

PERUSSE, M. 'Counting the near-misses' Occupational Health, 30(3): 123-126, 1978.

POWELL, P.I., HALE, M., MARTIN, J. and SIMON, M. '2000 Accidents: a Shop Floor Study of their Causes' London: National Institute of Industrial Psychology, 1971.

ROBAYE, F. 'Quelques propositions pour l'établissement d'un modèle de comportement dans les situations dangereuses' Bulletin du Centre d'Etudes et de Recherches Psychotechniques, 12(4): 331-340, 1963.

ROBAYE, F., HUBERT, J. and DECROLY, L. 'Estimation de la probabilité et de la gravité des accidents. Comparaison entre des groupes à haute et à basse fréquence d'accidents' Bulletin du Centre d'Etudes et de Recherches Psychotechniques, 12(4): 341-357, 1963.

ROSS, H.E. 'Behaviour and Perception in Strange Environments' London: George Allen & Unwin (Advances in Psychology Series), 1974.

RUMP, E.E. 'Cluster Analysis of Personal Questionnaires Compared With Principal Component Analysis' British Journal of Social and Clinical Psychology, 13: 283-292, 1974.

SCHROEDER, D.A. and LINDER, D.E. 'Effects of Actors' Causal Role, Outcome Severity and Knowledge of Prior Accidents upon Attribution of Responsibility' Journal of Experimental Social Psychology, 12: 340-356, 1976.

SELL, R.G. 'Ergonomics versus Accidents' Ergonomics for Industry No 5, Department of Scientific and Industrial Research (DSIR), 1964.

SLATER, P. 'Notes on Ingrid 72: A Program Available for Analysing Grids Individually under a Grant from the M.R.C.' London: Institute of Psychiatry, 1972.

SLOVIC, P. 'Convergent Validation of Risk-taking Measures' Journal of Abnormal and Social Psychology, 65(1): 68-71, 1962.

STARR, C. 'Social Benefit versus Technological Risk' Science, 165: 1232-1238, 1969.

SUCHMAN, E.A. 'Conceptual Analysis of the Accident Phenomenon' in 'Behavioural Approaches to Accident Research', New York: Association for the Aid of Crippled Children, 1960.

SURRY, J. 'Industrial Accident Research: A Human Engineering Appraisal' Toronto: Labour Safety Council of Ontario, 1969.

WALLACH, M.A. and KOGAN, N. 'Aspects of Judgement and Decision Making: Interrelationships and Changes with Age' Behavioral Science, 6: 23-36, 1961.

WALLACH, M.A. and KOGAN, N.'Risk Taking as a Function of the Situation, the Person and the Group' in 'New Directions in Psychology, Volume III (Part 3)'. New York: Holt, Rinehart and Winston, 1967.

WIGGLESWORTH, E.L. 'A Teaching Model of Injury Causation and a Grid for Selecting Countermeasures' Occupational Psychology, 46:69-78, 1972.

WILLIAMS, J.B. 'Knowledge and Attitudes in Relation to Toxic Substances: Benzene and Cyanide' MSc dissertation, Department of Safety and Hygiene, University of Aston in Birmingham, 1976.

WILSON, A. 'Attitudes of Workers to Hazardous Substances: Understanding and Evaluation of Hazards Due to Trichlorethylene, Asbestos and Ionising Radiation' MSc dissertation, Department of Safety and Hygiene, University of Aston in Birmingham, 1975.

WORKING PARTY 'The Acceptability of Risks' London: Council for Science and Society, 1977.