The Validity of Weighted Scoring Evaluation Techniques Applied to Design:

Studies in the Appraisal of Heating, Ventilating and Air Conditioning Systems for Office Buildings

A Thesis submitted for the Degree of Doctor of Philosophy by

Stuart John MacPherson

Department of Building Engineering and Surveying

Heriot-Watt University

May 1994

I believe that everything should be made as simple as possible

...but not simpler

Albert Einstein

Contents

List of Figures

List of Tables

Acknowledgements

It is with great pleasure that I record my thanks to those who have assisted me in this work. My supervisor, John Kelly, has provided support and guidance throughout the course of the research. Colleagues Roy Webb and Jack Torrance introduced me to the many senior building services consulting engineers who were essential to the empirical studies conducted during the research. Peter Aspinall advised me in the practical application of psychophysical measurement and repertory grid techniques. Stephen Vickers of Heriot-Watt and Val Belton of Strathclyde University allowed me to use their *VISA* computer programme and I am also grateful to them for the exchange of ideas on decision aids generally.

I also thank those researchers at other institutions who were (in some cases astonishingly) willing to engage in extended correspondence on a variety of topics related to the research. In particular, Benjamin Hobbs and J Robert Newman were a source of stimulating ideas regarding the validity of decision aids, and also assisted in keeping me up to date with current thinking by putting me in contact with other researchers, notably Norman Anderson, and providing me with advance copies of work which was unpublished at the time. I am also grateful to Nancy Johnson for valuable information on the teachback interviewing technique.

Finally, I thank my family for their encouragement and apologise to Jane whom I have shamefully neglected over the past two years during which this work has been concluded and written up.

Abstract

The common weighted scoring evaluation technique is presented by the design methods literature and by many practical guides as being an appropriate method for appraising different design solutions, however there are few critical assessments of the assumptions inherent in the method and no serious attempts to evaluate the validity of the technique as applied in a design context. This thesis presents a series of empirical studies and theoretical reviews which examine, in a logical sequence, aspects of the validity of weighted scoring techniques in the context of early stage heating, ventilating and air conditioning (HVAC) system design for office buildings. The nature of the HVAC design process is investigated, and in parallel with this a theoretical critique of the weighted scoring method as described in the design methods literature is conducted. It is found that the common approach to weighted scoring is invalid, raising concern over the indiscriminate use of such decision aids. However, a theoretically correct interpretation known as Multi-Attribute Value Theory (MAVT) is possible. It is also found that the method is not applicable to the selection of HVAC systems in general, but may be considered reasonably valid in more restricted tasks such as air conditioning system selection for a specific area in a building. While the MAVT models developed are judged to be reasonably valid, it is argued that their usefulness is debatable. If all the information on which to base the decision is available and the decision maker is reasonably skilled then MAVT will only improve decision making at the margin where the penalty for a wrong decision is less significant.

Chapter 1

Introduction

This chapter provides an introduction to the topic of the thesis and gives an overview of the research. The weighted scoring technique is briefly described and the perspective from which this decision aid is to be evaluated is outlined. The choice of domain, namely heating ventilating and air conditioning (HVAC) design in office buildings, is explained, and it is shown why this is a suitable starting point for a critical assessment of the validity of weighted scoring techniques in design generally. Finally, the structure of the thesis is outlined on a chapter by chapter basis.

1.1 BACKGROUND TO THE RESEARCH

The Royal Institution of British Architects (RIBA) plan of work (RIBA, 1973), in common with design guides for construction produced by equivalent institutions in other countries (Piper, 1987; Architektenkammer Nordrhein-Westfalen, 1986) and guides in other fields of design (eg VDI-Richtlinie 2222 Blatt 1 (Entwurf), 1973; Phal and Beitz, 1984), portrays design as a planned and methodical process. The design starts at a general conceptual level and then passes through a number of typical intermediate stages which are progressively more definitive to a final detailed design. At each stage, information on requirements for the building and its components is gathered and a range of alternative design solutions are synthesised at a level of detail commensurate with that stage. These solution options are appraised in terms of how they meet the various requirements and the 'best' solution is identified. This design solution is then broken down into smaller sub-problems and developed further at the next stage in the design.

The RIBA plan of work does not offer any specific guidance on how to perform an appraisal of different solutions, however the design methods literature does suggest that *weighted scoring* is an appropriate technique. This technique is also recommended as an additional assessment to the purely economic considerations of life cycle costing in Flanagan *et al* (1989), Kirk and Dell'Isola (1983) and in the Department of Health publication *Option Appraisal* (DHSS, 1987). More specifically, variations of the technique have been suggested as a decision aid for selecting among a range of HVAC solution types (Finch, 1989; Gilleard, 1988; Neuman and Guven, 1988).

1.1.1 Weighted Scoring Techniques

Consider a decision maker or *actor* (the term used in decision theory to mean one actively engaged in a decision process) who is faced with the task of deciding which option from among a range of possibilities is 'best' in a particular set of circumstances. Usually this will involve assessing the performance of each option on a range of criteria, and identifying the one with the superior overall performance. The actor may attempt to conduct this appraisal by examining the options on a purely qualitative level, noting their various advantages and disadvantages and then endeavouring to synthesise a judgement based on this. However, unless there are quite obvious differences between the options, a qualitative assessment of this kind may not be of much help. Moreover, if the number of items of information which the actor must balance in his mind exceeds about seven, the cognitive load imposed is great and would probably exceed his information processing capability (Miller, 1956; Silverman, 1991). It is well known that in these circumstances, the actor could resort to simplifying heuristics such as ignoring certain pieces of information, and so make selections which are sub-optimal (Montgomery, 1983; Fischer, 1979; Svenson, 1977).

In an effort to deal with these problems, it is often the practice to convert the qualitative decision problem into a quantitative one by using numbers to describe the performance of each option. The use of numbers as a decision aid in this way is not new, as the following extract from a letter written by Benjamin Franklin to Joseph Priestly in 1772 demonstrates (source Brownlow, 1989 p 49):

> "I cannot, for want of sufficient premises, advise you what to determine, but if you please I will tell you how...My way is to divide half a sheet of paper by a line into two columns; writing over the one Pro, and over the other Con. Then, during three or four days consideration, I put down under the different heads short hints of different motives that at different times occur to me for or against the measure. When I have thus got them all together in one view, I endeavour to estimate the respective weights...[to] find at length where the balance lies...And, though the weight of the reasons cannot be taken with the precision of algebraic quantities, yet, when each is thus considered, separately and comparatively, and the whole matter lies before me, I think I can judge better, and am less liable to make a rash step; and in fact I have found great advantage for this kind of equation in what might be called moral or prudential algebra."

In the field of design methods, which emerged as a distinct discipline in the 1960s, the influence of systems theory and management science led to the adoption of the quantitative method which is generally known as *weighted scoring* as a technique for the assessment of alternative design options. The common approach to this as described in the literature (see for example Jones, 1970; Sanoff, 1977; Phal and Beitz, 1984; ASCE, 1988; Cross, 1989) is as follows: First the criteria on which the options are to be appraised are listed. Next, the criteria are *weighted* with a number which reflects their relative importance (eg on a scale of 1 to 10). The options are similarly awarded a *score* on each criterion which reflects their performance. Finally the criterion scores are multiplied by the associated criterion weight and these weighted scores are added up to produce a total weighted score for each option. The option with the highest total score is then the one with the 'best' overall performance.

Variations on this technique also exist. In particular, some methods attempt to incorporate explicitly the decision actor's perception of uncertainty within the appraisal by adding to the elicitation of scores an assessment of the probability with which the performance score would be achieved (eg Moselhi and Deb, 1993). The underlying method is still, however, a weighted scoring technique.

While attractive, the weighted scoring technique as presented in the design related literature is an *a priory* method. It has reasonable face validity in that it seems to make sense, however there are few critical assessments of the assumptions inherent in the method and no serious attempts to evaluate the validity of the technique as applied in a design context. As it will be shown, wider reading of the literature on design theory and decision analysis suggests that there is good reason to question the validity of such techniques. This research aims to address this question of validity in the context of early stage HVAC design for office buildings.

1.1.2 The Early Stage HVAC Design Problem

According to the RIBA plan of work, the early stage engineering services design problem may be characterised as a choice between a range of generic system options. In particular, plan of work diagram 3 indicates that at the feasibility stage, it is the role of the services engineer to "state the feasibility of alternative services design solutions". Although no specific recommendation is made, it is implied that this appraisal will be carried out at a conceptual level of design with only the most basic form of drawn information, if any. This philosophy is reflected in the Building Services Research and Information Association (BSRIA) design manual for building services engineers (BSRIA, 1990) which follows the model of design laid down in the RIBA plan of work and enlarges on the function of the services engineer in the design team.

The selection of a generic solution is arguably the most important design action which the services engineer will take, as it has long been recognised that the implications for cost and quality are greatest for those decisions taken early in the design process (Ferry and Brandon, 1984). This thesis is concerned with the selection of generic heating, ventilating and air conditioning (HVAC) system options, with particular reference to office buildings. This particular domain was chosen because an inspection of published building analyses in the professional journals (eg *Building Services* and *The Architects Journal*) and of the HVAC design literature (eg CIBSE, 1986; ASHRAE, 1991; Legg, 1991; Stanford, 1988) indicated that it is in the servicing of offices where the greatest range of identifiably different generic HVAC solution types is to be found.

HVAC design is also a suitable starting point for a critical assessment of the validity of weighted scoring techniques in design generally as the early stage HVAC design task has been specifically characterised as the selection of a generic system type from among a range of options. Moreover, HVAC system design, while not being entirely devoid of aesthetic considerations, is dominated by economical and functional ones. The design methods literature does indicate that the assessment of options becomes progressively more difficult as aesthetic considerations become more important, as these are less amenable to quantification. The validity of weighted scoring techniques in the context of HVAC design may therefore be regarded as a minimal requirement, because if it is found to be invalid as a decision aid here, then there is little prospect of it being valid in other more general building design contexts.

1.2 ASSESSING THE VALIDITY OF WEIGHTED SCORING TECHNIQUES

There are two interrelated aspects to the evaluation of decision aids which have an important impact upon the methods which are used to perform the evaluation: these are the *perspective* from which the aid is to be evaluated and the *conceptual considerations* which are involved in arriving at a definition of validity.

1.2.1 Perspectives for the Evaluation of Decision Aids

Adelman and Donnell (1985) suggest that issues relevant to the evaluation of a decision aid arise at three interfaces which place different perspectives on the aid performance (see also Adelman, 1992). The interface between the decision aid and the user (actor) deals with the extent to which the aid is 'user friendly'. The second interface is between the user (and the decision aid) and the decision making organisation of which the user is a part, and this deals with the extent to which the decision aid assists the decision making of the organisation. Finally, the third interface, between the organisation and the environment, deals with the extent to which the decision aid improves the organisation's performance in the wider environment. A pictorial interpretation of this framework is presented in figure 1.1(a).

Riedel and Pitz (1986) suggest that this framework obscures some significant evaluation perspectives. They acknowledge that the Adelman-Donnell framework does highlight that when evaluating an operational decision aid it is often the aid-user-organisation combination that is actually being evaluated; and when evaluating the suitability of the decision aid for the organisation it is often the aid-user combination that is actually being evaluated. However, they go on to point out that this is not necessarily true. Most significantly, they assert that the primary perspective in evaluating a decision aid is the *validity* of the aid as applied to the *technical decision task*. If the decision aid is not valid in this sense, then there is little point in conducting further, more complex evaluations of aid-user-organisation combinations. This consideration is included in the organisation/environment interface of the Adelman-Donnell framework (the task existing in the environment), however Riedel and Pitz argue that its inclusion there obscures the importance of specifically technical, task related aspects which constitute the minimal requirements for validity. It is quite possible to assess certain aspects of the decision aid/task interface without involving the user or the organisation, or to assess certain aspects of the decision aid/organisation interface without involving the user. A modified framework based on these comments is illustrated in figure 1.1(b).

A proviso must be made, however, to the claim that certain aspects of the validity of the decision aid as applied to the technical task can exclude the user. This statement is true once a definition of the task exists. It should be noted, however, that in the context of decision aids, the definition of the task will normally originate from the user in the first place. Hence, the task cannot be considered as having some form of objective reality out in the environment, but rather will be the task as perceived by the user.

While it would be possible to conduct research at any of the interfaces illustrated in figure 1.1(b) this research is concerned with the fundamental issue of the validity of weighted scoring as a decision aid to the *technical task* of early stage HVAC design.

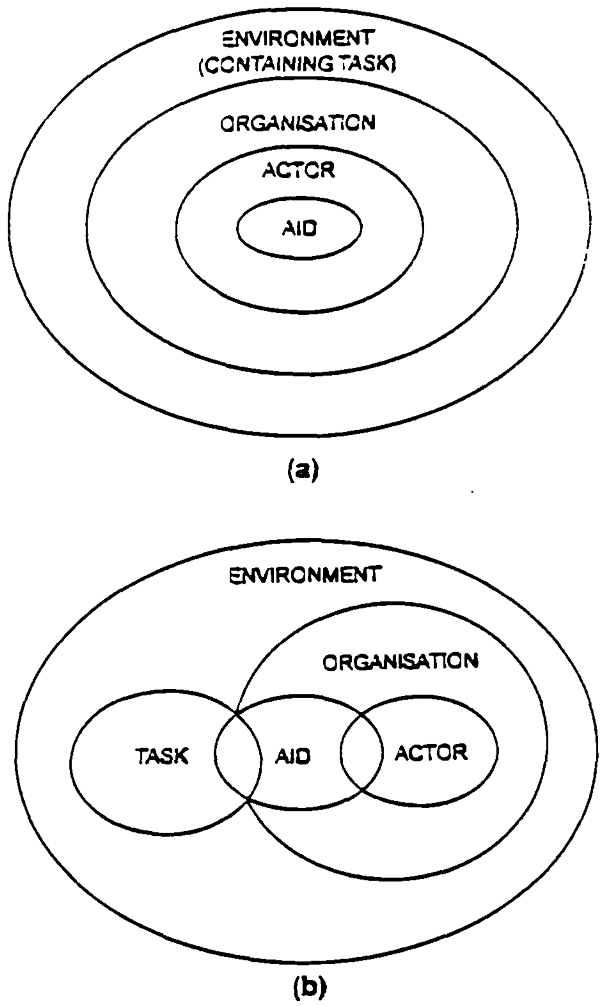


Figure 1.1 (a) A pictorial interpretation of the Adelman-Donnell framework

(b) A modified framework which accounts for the criticisms of Riedel and Pitz

1.2.2 Conceptual Considerations in the Evaluation of Decision Aids

The evaluation of decision aids is a complex conceptual problem to which there is as yet no comprehensive solution. This problem is discussed at several points in the thesis as it has profound implications for the assessment of the validity of weighted scoring techniques. The route of the difficulty is the fact that most decisions are value laden, in which case there can be no objectively `correct' outcome. Moreover, even if a correct outcome could be defined, it does not follow that the decision leading to the outcome was `correct'. The uncertainties which abound in realistic decision situations mean that good decisions can lead to bad outcomes and vice versa (Howard, 1988).

In the absence of an objective outcome based success criterion, decision theorists advocate a more eclectic approach to the validation of decision aids. The general principle behind this approach is that in the absence of a conclusive *outcome* based measure of validity, the quality of the decision *process* must be used as the measure of validity (Howard, 1988; Janis and Mann, 1977).

In the course of the thesis arguments will be developed which show that for the decision aid to be considered valid, it must be:

(i) compatible with the decision task (eg characteristics of the task should not be at odds with assumptions inherent in the decision aid)

(ii) theoretically consistent in its various constituent elements (eg in the definitions of criteria, weights and scores) and

(iii) should produce decisions which are defendable and are acknowledged by the actor as being superior to unaided judgement.

These considerations are in many respects *relative* and *subjective*. Consequently, decision aids cannot be validated by single critical experiments or disconfirmed in the usual Popperian sense. Rather, the goal of validation must be to construct a `web of evidence' consisting of different kinds of data, each of which addresses different aspects of the aid (Kuhn, 1970; Quine 1976).

1.3 RESEARCH METHOD AND OVERVIEW OF THE THESIS

It follows from the above conceptual considerations that the research methods adopted in this work differ from the conventional hypothesis - test methodology. This thesis presents a series of empirical studies and theoretical reviews which examine, in a logical sequence, aspects of the validity of weighted scoring techniques. Each study is important in its own right in that it uncovers something about the validity of weighted scoring in the context of HVAC design in office buildings, and also informs the subsequent stages of the research.

Firstly, to assess the compatibility of the weighted scoring technique with the HVAC design task it was necessary to understand the characteristics of the design task in general and HVAC design in particular. In chapter 2, the theories of design as a process are reviewed. It is shown that the model of design assumed by the weighted scoring technique and embodied within the RIBA plan of work and similar guides is 'systematic'. Since the early 1970s, however, the systematic model of design has come under considerable criticism from an empirical point of view, thus raising serious questions about the applicability of weighted scoring. These criticisms are discussed and arguments are put forward in support of the view that a systematic approach to design, supported by appropriate methods such as weighted scoring, is still a realistic and worthwhile goal, *provided the methods are adapted to the characteristics of the task to be performed.*

What characterises the early stage HVAC design task? According to the RIBA and BSRIA guides, it is typified by the appraisal of alternative generic HVAC solutions occurring at the feasibility stage. However following the review of chapter 2 it cannot be assumed that what is written in these publications is representative of how practising engineers perceive the technical design task. Should it be that the task is not characterised by an appraisal of different solutions, then then this fact might have dictated a different direction for the subsequent stages of this research.

Chapter 3 describes how a preliminary study was carried out to determine whether there is a 'typical' form of design description for HVAC systems at the feasibility stage, and to provide an understanding of the process by which the designs evolve. These research aims called for an inductive research method, whereby insights and explanations could be derived from data, rather than a positivist method which is concerned with prediction, measurement and the testing of hypotheses (Smith, 1981; Podaskoff and Dalton, 1987). Hence, this study employed a multiple case study research strategy in conjunction with a grounded theory approach to analysing the data gathered. It was found that while the design development appears on the surface to be very different to the top-down approach assumed by the systematic model, engineers do adopt a hierarchical process of HVAC design proceeding from the conceptual to the specific. This process of design seems to begin with a two stage appraisal in which the fundamental HVAC provision (eg natural ventilation, mechanical ventilation, air conditioning) for each area of the building is assessed and then an evaluation is made of generic system types from within this provision category. It is concluded that in the initial stages the 'typical' HVAC design process *can* be characterised as a decision between alternative solution types. *This does not automatically imply that weighted scoring is an appropriate method for carrying out such an appraisal*, however further investigation into the scope for the use of a formal decision aid such as the weighted scoring technique would seem justified in this context.

Before proceeding to study this decision making task in greater depth, it is first necessary to understand the theoretical background to decision making in general, and weighted scoring techniques in particular. Chapter 4 presents a review of the literature related to decision making and lays the philosophical basis for the continuing study. It is argued that while some writers suggest that decision making *per se* may have a limited role in solving problems in general, it has an important place in the typical HVAC design process. Different models of the decision making process are discussed and it is shown that, as with design theory, there is a difference between the prescriptions of the systematic model of decision making and how decisions are made in practice. The weighted scoring technique is placed in context within the systematic model as a prescriptive appraisal technique. The rational behind prescriptive appraisal techniques is discussed and weighted scoring is compared with the well known appraisal technique of cost-benefit analysis. Finally, the vexing question of how to assess the validity of a technique, which has implications for the research methodology, is considered once again.

In chapter 5, a detailed theoretical critique of the weighted scoring technique as it is normally described in the design methods literature is presented and it is demonstrated that by the standards of mainstream decision analysis this simple version of the technique is theoretically invalid, raising concern over the indiscriminate application of such techniques. It is shown that a theoretically correct interpretation of the weighted scoring technique known as Multi Attribute Value Theory (MAVT) is possible, however it is pointed out that the validity of MAVT is in some ways a matter of degree rather than principle. In the course of this theoretical review, it becomes clear that MAVT makes certain important assumptions about the decision task which means that *it is not a valid technique to apply in every situation that may be characterised as an appraisal of different solutions*. Having reached a clear understanding of these limitations, it was necessary to conduct a further study of the solution appraisal task in early stage HVAC design which was identified in the preliminary study of chapter 3 to ascertain the extent to which this task is amenable to the application of MAVT.

Chapter 6 relates how a second empirical study was carried out to gain a greater understanding of the characteristics of the HVAC solution appraisal task. Once again this called for an inductive research method whereby a model of the task characteristics could be constructed from data. The study used techniques developed originally in the field of expert knowledge acquisition to build a partial model of the early stage technical solution appraisal task as perceived by consulting engineers. The results of this study indicated that the task of deciding upon the fundamental HVAC provision for a space (eg natural ventilation, mechanical ventilation, air conditioning) is *not* amenable to the application of methods like MAVT, but it was found that the more specific task of selecting a system type within a major provision category may be an appropriate problem for MAVT, with further research being most usefully concentrated on air conditioning system selection.

This then sets the scene for the third and largest of the empirical studies in this research. Chapter 7 addresses the methodological difficulties associated with assessing the validity of a decision aid which were introduced in chapter 4, a problem to which there is no pat answer, and describes the methods adopted here which it is hoped take some steps towards making an assessment of validity possible. Chapter 8 relates the application of these methods to a study in which six subject engineers are presented with three hypothetical case vignettes and use MAVT to select the most suitable generic air conditioning system type from among a range of options for each case. It is concluded from this study that *it is not possible to develop an MAVT model which is a complete aid to the appraisal task* because some aspects of the task violate the assumptions inherent in the technique. Models were finally constructed which had to compromise by omitting certain criteria, rendering some decision considerations external to the model. Nevertheless, the MAVT models are substantially if not totally complete, and so *within this limitation it was found that MAVT can be considered a valid technique when applied to this task.*

Finally, chapter 9 summarises the results and conclusions of the whole of this research, examines some practical issues concerning the application of MAVT which have emerged, and comments on areas which are considered important for further study. In particular, it is highlighted that while the evidence is that MAVT is valid and did produce superior decisions to unaided judgement, these improvements were evident only in marginal instances where no serious penalties would be incurred in selecting the next best alternative. Hence, although the MAVT technique may be valid, there must be some question over its *usefulness*.

It is believed that this problem arises dilemmatically from the rigorous validity requirements of MAVT, and so the relaxation of these requirements coupled with applications of the technique in real design settings should be considered a future research priority.

Chapter 2

A Review of Design Theory

Over the past 30 years, there have been considerable developments in the theory of design as a process. This chapter presents a review of these developments, in particular, in relation to the orderly approach to design assumed by weighted scoring methods. The current models of design embodied within the RIBA plan of work and similar design guides are examined and it is shown that these have a common approach which has been classed as 'systematic'. Since the early 1970s, however, the systematic model of design has come under considerable criticism from an empirical point of view. These criticisms are discussed and arguments are put forward in support of the view that a systematic approach to design, aided by appropriate methods, is still a realistic and worthwhile goal, provided the interpretation of such an approach and its methods is sufficiently flexible.

2.1 THE SYSTEMATIC DESIGN MODEL AND WHY IT IS CRITICISED

The RIBA plan of work (RIBA, 1973), in common with design guides for construction produced by equivalent institutions in other countries (Piper, 1987; Architektenkammer Nordrhein-Westfalen, 1986) assumes a planned and methodical approach to design development. The design starts at a general conceptual level and then passes through a number of typical intermediate stages which are progressively more definitive to a final detailed design. At each stage, information on requirements for the building and its components is gathered and design solutions are synthesised at a level of detail commensurate with that stage. These solutions are then appraised in terms of how they meet the requirements prior to being broken down into smaller sub-problems and developed further at the next stage in the design. This is also the view of design adopted by guides in the field of engineering design (eg Phal and Beitz, 1984; VDI-Richtlinie 2222 Blatt 1, 1973).

2.1.1 The Two Dimensional Systematic Design Model

Markus (1969) and Maver (1970) formalised a model of architectural design which portrayed the underlying procedures in these design guides in two dimensions as shown in figure 2.1 overleaf. This representation of the design process is a consolidation of several similar models which were developed in the 1960s by various authors in different fields of design (eg Asimow, 1962; Jones and Thornley, 1963; Archer, 1965; Luckman, 1967) and has its origins in the general systems engineering approach to the design of complex technical or social systems which can be traced back to Hall (1962).

In this systematic view, the vertical dimension corresponds to the development of the design, while the horizontal dimension corresponds to the problem solving/decision making activity which occurs at every stage of development. In this horizontal problem solving/decision making activity, an iterative process of optimisation occurs as various design solutions for each problem (or sub-problem) are synthesised and appraised in an effort to find the 'best' one available.

During the 1960s, researchers in design theory developed a number of design methods for use within the framework of the systematic design model. These methods ranged from ways of setting design requirements to ways of selecting the best design solution from a number of alternatives (such as weighted scoring). Several of these methods have been collected in Jones (1970) and will be discussed further in section 2.2.4.

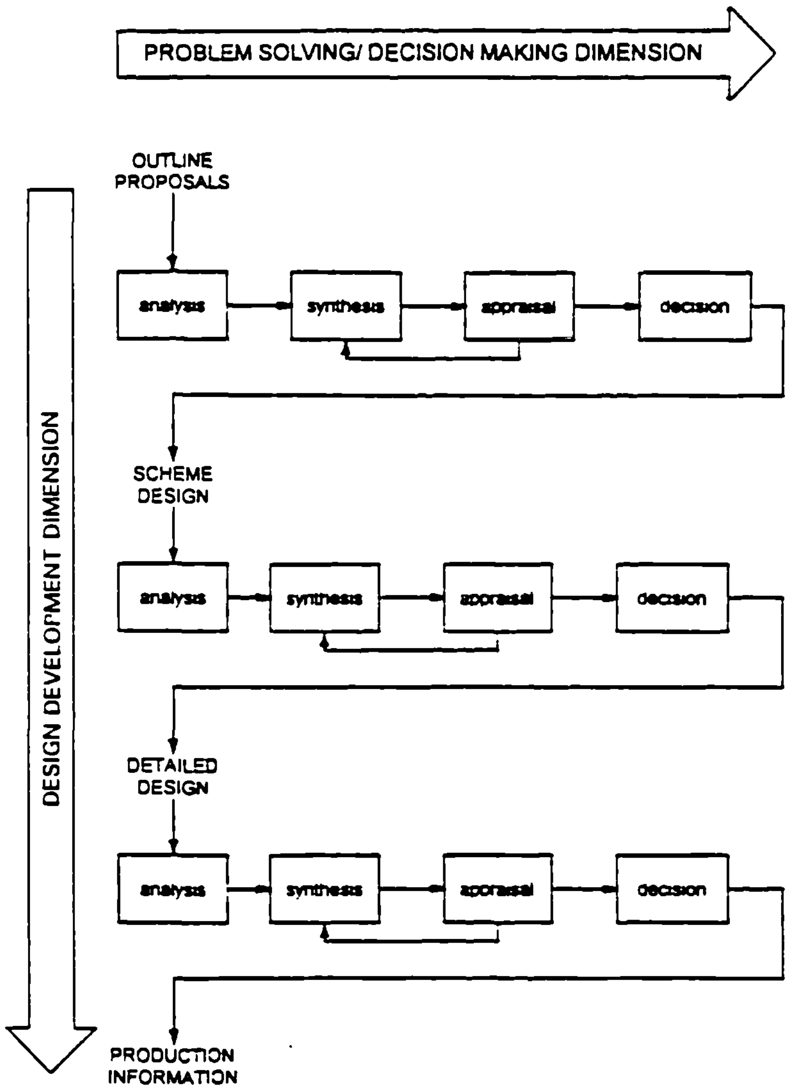


Figure 2.1 The Markus and Maver model of architectural design

In the 1970s, however, criticism of the systematic view of architectural design as exemplified by the Markus and Maver model began to emerge, and more recently similar criticisms have been made of the systematic model in the field of engineering design. These criticisms have simultaneous implications for both the horizontal and vertical dimensions of the model, however in the interests of clarity the main points of the arguments will be treated separately here in the context of each dimension.

2.1.2 Criticisms of the Horizontal Dimension

In relation to the problem solving/decision making dimension, Hillier *et al* (1972) argued that the analysis-synthesis process whereby a thorough analysis of stated requirements should be carried out before attempting to synthesise a solution from this data, is based on a fallacious and out-dated belief in the use of pure inductive logic in science. They suggested that design is only possible in light of a prior knowledge of solution types, and in fact the process of design is one in which the designer will first conjecture a solution and then subject this to analysis and appraisal. Eastman (1970) and Lawson (1980) observed designers at work in laboratory settings and found that designers do not attempt an exhaustive analysis of the available data before attempting to synthesise a solution, but rather that they make early attempts at solutions as a way of trying to understand the problem more fully. Darke (1979) conducted interviews with architects charged with the task of designing local authority housing and found that in an effort to understand the various complex aspects of the problem, they would propose a form of solution at an early stage and then use it to further explore the problem and highlight areas in which they had insufficient information. As a result of this research she developed the partial model of the design process shown in figure 2.2.

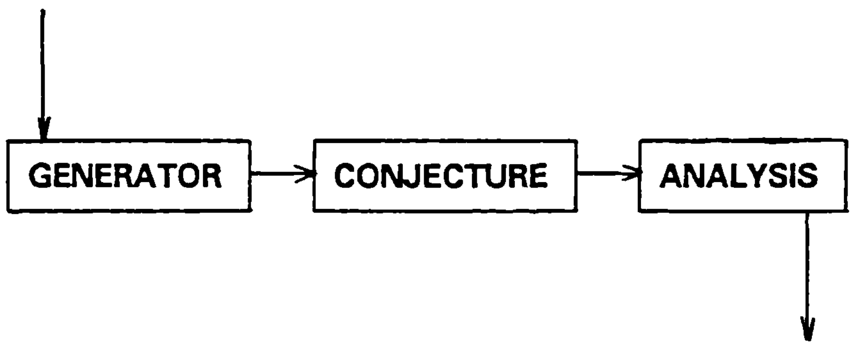


Figure 2.2 Darke's partial model of design

In a similar vein, Archer (1979) suggested that the actual practice of design is commutative where the designer's attention oscillates between emerging requirements on the one hand and the emerging solutions and sub-solutions on the other. The requirements influence the solutions and the solutions influence the perception of the requirements as the designer progressively reduces the misfits between them.

It would seem that the design research community is generally agreed that this paradigm is closer to a description of how design is actually performed in the horizontal dimension than any other model yet proposed (Willem, 1991; Roozenburg and Cross, 1991).

2.1.3 Criticisms of the Vertical Dimension

A consequence of the oscillating nature of the problem solving/decision making activity described above is that solutions and sub-solutions develop in a somewhat unpredictable fashion. An initial conjectured solution to one part of the design must be progressively refined and checked for compatibility with other emerging aspects of the design. Consequently, the designers find themselves operating at a variety of levels of detail simultaneously, which is at odds with the 'top-down' assumption of the systematic design model.

Early reservations on the ability to hierarchically de-compose a design problem into a stable structure of sub-problems and then proceed to solve them in a top-down fashion were voiced by Alexander (1966). Lawson (1980) noted that the model of architectural design produced by Markus and Maver (figure 2.1) indicated no feedback between the vertical stages, when in practice every stage has to be accessible to feedback from any subsequent stage. He argued that the emergence of some design details in one sub-system of the building can have a retro-acting effect on the higher level design of that system or of another, interdependent sub-system. A similar point was made by Hickling (1982). He suggested that the linear (horizontal) notion of iteration is incorrect, and that a cyclic 'whirling' process moving unpredictably in both dimensions of the systematic model is more appropriate. Mackinder and Marvin (1982) observed that even at the early stages of design, architects could be observed to stop working on the scheme design of a building to work out a particular detail which they considered critical to the feasibility of the larger design. This suggests a form of 'feedforward' activity between vertical stages in addition to the feedback discussed by Lawson and Hickling. Ullman *et al* (1988) and Takeda *et al* (1990) have also made similar observations of the engineering design process. They concluded that contrary to the systematic design model, the design process is not uniformly hierarchical, proceeding from the conceptual level towards detailed design. Instead, designers shift their attention from one aspect of the design to another in an opportunistic manner and so details develop in an unpredictable, fragmentary fashion.

2.1.4 The Prescriptive/Descriptive Dichotomy

Commenting upon these criticisms of the systematic design model, Cross (1989) points out that, in retrospect, it can be seen that the attention of the researchers in this field has shifted from what designers *aught* to do, to what designers are *observed* to do. By emphasising the sequence of stages through which a design should progress, the systematic model of design (and any guide based upon it) is *prescriptive*. It is intended to offer guidance on more effective methods of working. On the other hand, the models which were developed based on observations of actual design activity such as the one produced by Darke (figure 2.2) are *descriptive*. They seek to describe the actual thought processes and conduct of designers.

The fact that the models are different does not, *per se*, invalidate the systematic model or design guides, and criticisms to this effect are over-stated. For example, Kraus and Mayer (1979) suggested that "if any conceivable strategy, list of operations or route is permissable in finding a solution, then none can be prescribed as mandatory". This is certainly true, but it does not disqualify models which are intended merely to structure and not predict the course of design activities. Taking models at face value in this way and thereby misinterpreting them is termed 'model blindness' by Powell and Russell (1983). They point out that underlying every model of something as complex as the design process lie a set of simplifications and assumptions of which the model builders are well aware. There is however a danger of misinterpretation when users and critics of the models bring to bear their own set of assumptions and brand of common sense.

As Roozenburg and Cross (1991) put it, the systematic design model is a weak model, because following it "requires 'sensible' (knowledgeable/informed) interpretation by the designer of the vaguely defined 'rules' and terms, and, even if properly applied, success is not guaranteed". Such a model, they say, is not intended to provide a recipe, but rather tries to "organise the problem solving behaviour of designers to such an extent that it is more effective and efficient than intuitive, unaided, unsystematic ways of working".

2.2 LIMITATIONS OF DESIGNERS AND LIMITATIONS OF METHODS

It has been stated that the difference between the models of the design process which evolved in the 1960s and 1970s reflect a prescriptive/descriptive dichotomy in the underlying purpose of the models. This (in retrospect simple) taxonomy resolves the tensions which are evident in some of the literature of the 1970s concerning the difference in the models, and exonerates the systematic model, which was often accused of being 'wrong' because designers were not observed to follow it in practice. However, the taxonomy does not in itself explain why design practice differs from the prescriptions of the systematic model. In this section the limitations of designers and the limitations of methods when faced with complex design problems will be discussed, and the reason for the difference between prescriptive and descriptive models thus explained.

2.2.1 The Ill-Structured Nature of Design Problems

The most frequently sited characteristic of general design problems is that they are ill-structured (Willem, 1991). As Simon (1973) points out, 'ill-structured problem' is a residual concept, because it is only possible to demonstrate that a problem is ill-structured by showing that it is *not* well-structured (just as it is only possible to classify something as a UFO by showing that it is *not* identifiable in terms of known aerial phenomena). Several writers have, from time to time, set down criteria by which it is possible to define a well-structured problem, and in so doing have shown that general design problems are ill-structured because they fail to satisfy these criteria. Principle among these criteria, as summarised by Rittel and Webber (1973), are that design problems have no definitive formulation, have a more than countable number of interdependent sub-tasks and have no objective criterion for solution. They suggested that such ill-structured (or 'wicked' problems as they termed them) can never be 'solved', but merely 'resolved', and the acceptability of any given solution will depend simply upon how it compares with other possible solutions in the judgement of the designer.

It is precisely the ill-structured nature of the problems with which a designer must cope that makes the systematic design model seem impossible to follow. In the horizontal (problem solving/decision making) dimension, he is advised to follow a systematic process of analysing all the requirements for the design, then synthesising a solution or solutions from those requirements and appraising these solutions in terms of how they satisfy the requirements until an optimal design has been arrived at. However, to take this procedure *literally* is not only impractical, but impossible because of the infinite number of options that would need to be considered. Firstly, the synthesis of solutions would require the designer to consider every kind of material (brick, steel, concrete, wood, yak skin); in conjunction with every kind of morphology; in conjunction with every conceivable way of breaking the problem down into sub-problems. Secondly, there can be no initial definitive criterion by which a solution can be judged, because such a criterion can only become operational in the context of the solutions and sub-solutions produced (which cannot be predicted in advance) (Luckman, 1967).

2.2.2 How Designers Cope with Ill-Structured Problems

Simon (1973) suggests that when confronted by a complex design task, such as designing a building, the designer will deal with it by breaking it down into a set of smaller sub-problems which may be temporarily treated as well-structured. It is acknowledged that the way in which these sub-problems are delineated, and the order in which their solution is tackled will very much influence the final, overall design solution.

Simon goes on to say that, of course, the danger in solving problems in this manner is that interrelations between the various sub-problems are likely to be under-emphasized, however in practice these dangers are mitigated by the designer's skill in managing his approach to the design. He learns from training and experience that certain ways of dividing the task will make for smoother co-ordination of the sub-problems than others, and in this way, the solution of complex problems is made tractable.

Powell et al (1984) observed that when attempting to deal with complexity, designers adopt what they termed an 'information rejection strategy' where, to avoid information overload, the designer makes some fundamental decisions (eg building plan shape) early on to reduce the vast range of possible design solutions he would otherwise need to consider. This finding is supported by Mackinder and Marvin (1982) and Newland et al (1987), who reported that early scheme designs were drawn up by a senior architect with little reference to any written information (suggesting heavy reliance on his own experience). They further observed that there was almost never any serious consideration given to alternative designs, mainly due to shortage of time, and so final detailed designs usually followed the first attempt at an outline design very closely. Simon (1975) and later Maver (1979) also suggested that 'style' in design is in fact a mechanism for rejecting a large range of possible solutions and thereby reducing the design problem to more manageable proportions.

2.2.3 Propensity for Sub-Optimisation in Designers' Behaviour

This strategy of rejecting information and rapidly reducing the range of solutions which might be considered is not without its problems, however. The danger is that designers might overlook better design alternatives, and by relying so much on experience of similar building types, can be guilty of perpetuating the use of inferior designs. It is this tendency for designers to rely so much on preconceptions which led to the advocation of the systematic model by theorists in the first instance (Hall, 1962). The aim of the model was to encourage designers to reflect on the design requirements more thoroughly, and to consider a wider range of potential solutions in an attempt to arrive at a more optimal solution.

2.2.4 Limitations of the Systematic Approach

As pointed out in section 2.2.1, a literal interpretation of the systematic approach is unworkable in the context of ill-structured problems. Nevertheless, the propensity for sub-optimisation in design tasks is obvious, and hence the attempts to encourage designers to adopt a more systematic approach. During the 1960s, researchers in design theory developed a number of design methods for use within the framework of the systematic design model. Several of these methods have been compiled in one volume by Jones (1970). By way of illustration, some examples of these methods and how they apply within the systematic design model are given in table 2.1.

ANALYSIS	• Objective Trees • Interaction Nets • Interaction Matrices • Alexander's Method of Determining Components
SYNTHESIS	• Brainstorming • Synectics • Page's Cumulative Strategy • Collaborative Strategy for Adaptable Architecture
APPRAISAL	• Weighted Scoring • Checklists

Table 2.1 Examples of design methods

Jones, in reflecting upon these early design methods (Jones, 1977), states that a minority of them were used with mixed success, while most of them never found their way into wider use. It appears that the major failing of many of the methods lay in underestimating the difficulties associated with ill-structured problems. Any attempt to force the design problem into a rigid, once-and-for-all pre-determined structure was doomed, hence the substantial failure of methods such as Alexander's Method of Determining Components, Page's Cumulative Strategy and Collaborative Strategy for Adaptable Architecture.

Considering the few methods which met with some success, it would seem that a key element is the ability to apply them, at the discretion of the designer, to problems which are more vaguely defined, and may have only a temporary structure. "Good methods are built upon rationality, adapted to the characteristics of the task to be performed, and to the cognitive properties of the designer. This calls for an integration of prescriptive and descriptive insights" (Roozenburg and Cross, 1991).

This more pragmatic and flexible view of the design methods echoes the spirit in which the systematic design model itself must be applied. However sub-optimal actual design practices appear to be, it is now realised that many aspects of the designer's skill (and indeed the skill of humans in general) in dealing with complex, ill-structured problems are truly awe inspiring, and will defy any foreseeable attempt at complete systemisation (Beheshti, 1993).

2.3 THE EFFECTIVENESS OF THE SYSTEMATIC APPROACH

In section 2.1.4 it was stated that the systematic design model is a weak model because, as Roozenburg and Cross (1991) put it, following it "requires 'sensible' (knowledgeable/informed) interpretation by the designer of the vaguely defined 'rules' and terms, and, even if properly applied, success is not guaranteed". Moreover, success is itself an elusive concept because in the context of ill-structured problems there can be no objective criterion for solution (section 2.2.1). It has also been said above that the same comments apply to the design methods which may be employed within the framework of the systematic design model. In the absence of any success criterion, how do we know that the application of a process or method makes any improvement over unaided ways of working?

In the same paper, Roozenburg and Cross say that "there is until now little empirical evidence neither in favour of the effectiveness of the [systematic] model in contrast with conventional ways of working, nor against it". Consequently, they say that the value of the systematic model is "still largely based on personal experiences and belief in its rationality".

The difficulties associated with the testability and utility of prescriptive models or methods will be examined more fully in chapter 4, however the main issues are outlined below.

2.3.1 The Testability of the Systematic Model and Methods

The implication by Roozenburg and Cross that the effectiveness of the systematic design model (or methods it might be added) is open to empirical test requires some examination. It has been demonstrated from a philosophical point of view by Popper (1963) that methodologies (such as the systematic design model and design methods) are not open to empirical refutation. This is because they have no necessary correspondence with an objective reality (ie in the way that a scientific theory must correspond with experimental observation). This principle has since been expanded upon in relation to planning and design by Los (1981) and Faludi (1986). To return to the prescriptive/descriptive dichotomy, it could be said that a descriptive model of design is empirically refutable because to be correct it must correspond to the observed behaviour of designers. However a prescriptive model is not refutable in the same way, because there is no such requirement for correspondence. This is a more formal way of explaining why the fact that descriptive models of the design process differ from the prescriptive model does not invalidate the prescriptive model (section 2.1.4).

A prescriptive model may be tested theoretically by logical argument. It must be demonstrable that the model or method has a sound, rational basis. Empirical tests must still have a vital role however, because no matter how sound and rational the model may be theoretically, there is no point in prescribing actions which are impossibly difficult for designers to carry out, or too trivial to be useful in real situations. Because the principle of refutation does not apply, the validity of a prescriptive model cannot be confirmed or disconfirmed by single critical studies. Rather, the goal of model validation must be to construct a 'web of evidence' consisting of different kinds of data, each of which addresses different aspects of the model (Kuhn, 1970; Quine, 1976).

2.3.2 The Utility of the Systematic Model and Methods

It has been argued by some critics that the systematic model of design must be treated as such a weak model that it ceases to be useful. For example, in the field of planning theory, Scott and Roweis (1977) commented that "assertions that the planning process should involve various forms of goal formulation, evaluation, implementation and control, learning, and all the rest are no doubt true, but they are also plainly vacuous."

This criticism is understandable because although the systematic model is prescriptive, it merely directs designers in terms of *what* to do without informing them *how* it can be done. However, the systematic model is not vacuous in the sense that it provides a statement of the philosophy of the systematic approach to design (the *what*) which must be agreed upon before any progress can be made in developing methods to facilitate the approach (the *how*). It is clear, however, that without methods the systematic model is of limited use.

The design methods themselves must compromise between being on the one hand overly rigid and thereby too demanding to apply or simply incompatible with the ill-structured nature of design problems, and on the other hand being too weak to offer useful guidance. This is a difficult compromise to strike, however the continued advocation of certain methods such as weighted scoring suggests that designers believe that some methods do have a practical application. Many prominent design theorists too (like Roozenburg and Cross who have been sited extensively here) stick stalwartly to the view that some methods, correctly used, do offer an improvement over unaided ways of working.

In the next chapter, the first steps will be taken in an investigation of the early stage HVAC design process in an effort to discover whether the there is any scope for the use of the weighted scoring evaluation method.

Chapter 3

A Preliminary Study of Early Stage HVAC Design Development

The preceding chapter has discussed the differences observed between the prescriptions of the systematic design model and the actual behaviour of designers. The picture of actual design practice which emerges is a somewhat haphazard one which is highly idiosyncratic. It therefore seems very unlikely that the development of HVAC designs in practice follows the neat, top-down structure of the RIBA plan of work and the congruous BSRIA guide for engineering services design referred to in section 1.1.

As was argued, this does not preclude the development of more systematic methods of working, but to be useful these must be "...adapted to the characteristics of the task to be performed and to the cognitive properties of the designer. This calls for an integration of prescriptive and descriptive insights" (Roozenburg and Cross, 1991). To know if there is any scope for the use of the weighted scoring evaluation method described briefly in chapter 1, it is first necessary to know more about the design task as perceived by engineers, particularly in relation to office buildings, which will be manifested in the evolution of HVAC designs as it occurs in practice.

To recap, section 1.1 related that according to the RIBA and BSRIA guides the major consideration of competing alternative HVAC designs should occur at the feasibility stage. Although no specific recommendation is made, it is implied that this will be done at a conceptual level of design with only the most basic form of drawn information, if any. This chapter describes how a preliminary study was carried out to determine whether there is a 'typical' form of design description for HVAC systems at the feasibility stage, and to provide insights into the process by which the designs evolve.

3.1 THE RESEARCH METHOD FOR THE PRELIMINARY STUDY

The research aims outlined above called for a research strategy which can gather data which is sufficiently rich to provide essentially qualitative insights into design approach, and which is adequately protected against reporting biases by respondents. For example, if simply asked 'do you carry out an appraisal of a range of HVAC design options before deciding upon which system type to use?' it is highly likely that a design engineer will answer in the affirmative, because he wishes to appear conscientious. Moreover, the strategy had to be a flexible one, because due to the exploratory nature of the research, the 'research questions' were not well formulated initially, and would only take shape as information was gathered.

Further to this, it was also necessary to generalise to an extent on the results of the study, that is to be able to control for highly individualistic practices used by any one engineer and so make some inferences about the approach of engineers *in general* to the design of HVAC systems at the feasibility stage according to the RIBA plan of work. For the strict purposes of this research, this generalisation did not have to go beyond office buildings, however it was considered that it would be informative to extend the study into other building types.

Issues of methodological validity (such as reporting bias and generalisation) have been confronted by many social scientists, and a formal statement of four validity criteria which address these issues has emerged. These are namely, construct validity, internal validity, external validity and reliability (Yin, 1989). *Construct validity* involves establishing correct definitions and operational measures for the concepts being studied. *Internal validity* requires that where causal relationships are to be inferred from the data gathered, it can be demonstrated that reasonable steps have been taken to verify that these relationships are not spurious. *External validity* means establishing the extent to which the results of the study are generalisable. And finally, *reliability* requires that the operations of the study (ie data gathering and analysis) can be repeated with the same results.

In view of these methodological requirements and research objectives, it was decided to adopt a *multiple case study* research strategy to gather the information, while using a *grounded theory* approach to analyse the data.

3.1.1 Information Gathering using Case Studies

The use of case studies is one of the favoured information gathering research methods in the study of design (Bruce, 1993). Advocates of the method believe that case studies permit understanding of a phenomena, with meaning and authenticity. "Case studies are the preferred strategy when *how* and *why* questions are being posed..." (Yin, 1989 p11). The case study approach is rooted in an inductive tradition whereby insights and explanations are derived from data, which is distinct from the positivist approach which is concerned with prediction, measurement and the testing of hypotheses. Langrish (1993) and Svengren (1993) maintain that the method yields a richness and quality of data which is generally lacking in the positivist research model.

The Multiple Case Study Research Strategy

The use of *multiple* cases enhances external validity by allowing the inferences drawn from one case to be checked against the others. If the findings are found to match, they are therefore more likely to be valid in a general sense. It should be noted that in case study research, the objective is to generalise the results to theoretical propositions and not to populations. In this sense they are more analogous to experiments, which also aim at analytical generalisation, than surveys which aim at statistical generalisation and therefore must employ representative sampling techniques (Yin, 1989).

In this research, data was collected firstly by detailed inspection of the project files, and in particular that section of the files which the project engineer had flagged as constituting the feasibility stage. The files contained, in (usually reverse) chronological order, all of the important information on each job which had been committed to paper such as correspondence with the client and other consultants, minutes of meetings, internal memos, design calculations and drawings. Following this, an unstructured interview was held with the project engineer.

The inspection of the files was made initially with a view to obtaining a 'snap-shot' of the design at the conclusion of the feasibility stage. By then going back into the file and questioning the project engineer, it was possible to construct explanations (Yin, 1989) for the way in which the design had developed. In an effort to ensure the reliability of the data, the questions concentrated on facts and events, rather than highly subjective interpretations, and used standard inquisitorial courtroom interrogation (eg What did you do? When? Why? Who said what to whom? (Bourgeois and Eisenhardt, 1988)).

By using more than one source of evidence (project documentation and the recollections of the project engineer) it was possible to improve the construct validity of the research and achieve 'triangulation' (Smith, 1981). That is to say, that while there are obvious and well documented disadvantages in relying solely on the recollections of an individual (Nisbett and Wilson, 1977; Watson and Evans, 1975) and while there are equally disadvantages in relying on various items of information in files which may have failed to capture more subtle aspects of the design process, the requirement for corroboration on important issues between the data sources significantly enhances the dependability of the information.

The Cases Studied

Four consulting practices supplied information on a total of eight building projects from their current or very recent portfolio of work. Of the four practices, two were multi-disciplinary (incorporating architectural, services and structural engineering disciplines), one was both a services and structural consultancy and one was specifically a services consultancy. All of the practices were capable of dealing with large projects with a substantial HVAC content.

At the request of the researcher, the projects were selected by the HVAC project engineer to represent 'typical' examples of building and HVAC designs at the feasibility stage according to the RIBA plan of work. ('Typical' was taken to mean that the projects were not especially *atypical* in any way). As it was envisaged that the later stages of this research project would focus on office buildings, four of the buildings selected were offices of varying sizes and one additional project incorporated a substantial office provision. Three other building types were also investigated, however, to act as a check on the generality of the findings across different building types.

Table 3.1 summarises the development of the design description for each project at the conclusion of the feasibility stage and appendix 1 contains a fuller narrative which describes the projects in greater depth. Space and the need to respect confidentiality prevents the inclusion in the thesis of detailed information on each case project. However, an attempt has been made to include sufficient data to impart a 'flavour' of the project and to highlight the main points of evidence which provided justification for the findings of the study presented in section 3.2

Project number	1	2	3	4	5	6	7	8
Description	Office of 3500 sq m	Office of 8000 sq m	Office of 50 000 sq m	Office of 8000 sq m	Factory of 6000 sq m	Printing works of 1000 sq m	Laboratory of 1800 sq m	Library annexe of 1500 sq m
Architectural and structural								
Plans	General plans 1:200 Audio-visual room 1:20	General plans 1:200	General plans 1:100	General plans 1:100	General plans 1:200	General plans 1:100	General plans 1:100	General plans 1:100
Gross space allocation	All areas	All areas	All areas	All areas	All areas	All areas	All areas	All areas
Specific layout development	Open-plan office spaces Audio-visual room Main reception	Office areas	Office areas Circulation spaces	Office spaces	Office spaces Some production spaces	Storage areas Production areas	Entrance and exits Sterile laboratories Secure plant rooms	Entrance Service desk Circulation space
Structural grid and member sizes estimated	Office areas	Office areas	Floors above ground	Generally	Office spaces	None	Generally	Generally
HVAC services								
Generic system type decided or choice narrowed down	All areas	All areas	All areas	All areas	All areas	All areas	All areas	All areas
More detailed system design	Open-plan office areas Audio-visual room	Open-plan office areas	Open-plan office areas Ice storage system Roof located plant	Open-plan office areas	Modular office space Main services to factory	Certain production areas	Highly serviced labs Fume cupboard exhausts	Open-plan areas Roof located plant
Sections	Open-plan areas 1:20	Open-plan areas 1:20	None	Open-plan areas 1:20	Modular office space 1:50 Main factory services 1:50	As above 1:20	Labs as above 1:20	Open-plan space 1:20

Table 3.1 Summary of the development of the design description for each project

3.1.2 Analysis using Grounded Theory

Being rooted in an inductive tradition, case study research is not approached form a classical hypothesis-test methodology. Instead, the data from each case must be examined and a theoretical account of what happened and why allowed to emerge from the data itself. These theoretical accounts must then be compared across all of the cases to see whether any consistent pattern can be decerned. Precisely how this is done in the field of design research is generally left to the ingenuity of the individual researcher:

> "Having completed several cases, how does one 'add them up' to reach conclusions? This is where skill, imagination, lateral thinking and perseverance come in. There are no physics-type rules on how to identify new classes in a taxonomy or how to create new labels. Computers can help...but at the end, a bit of creativity is needed, its rather like design." (Langrish, 1993).

Although no recipe exists for the analysis of case study data, there is nevertheless an approach termed *grounded theory* (Glaser and Strauss, 1967) which offers a strategy for analysis.

This strategy for analysing data, originally developed in the social sciences, is termed the grounded theory approach because the researcher develops his own local theory to explain his particular observations which is *grounded* in the data rather than imposed externally from a general theory. The approach is considered appropriate when it is impractical to develop operational hypotheses from general theory either because there is no existing theory or because the theory is too abstract and remote to offer guidance in a specific case, which is precisely the position with design theory. The theoretical background provided by the design research literature described in chapter 2 created some expectations as to what would be found, for example the variation in design detail. However, such ideas could not provide operational hypotheses. They could be treated only as 'sensitising concepts' (Patton, 1980), providing an indication of what to look out for as the study commenced.

In this study, the grounded theory strategy also enhanced the internal validity of the research findings because the patterns decerned could be matched across all of the cases (Yin, 1989) and because the patterns themselves emerged from the data, rather than being imposed on the basis of an existing theory. It is when a compelling theory is being empirically examined using the conventional hypothesis-test research approach that there is a danger of preconceptions shaping the researcher's perception of the data (Svengren, 1993).

On a very practical level, grounded theory is also well suited as a method for dealing with qualitative data of the kind gathered from case studies involving unstructured interviews and the examination of documentary sources (Turner, 1981). As Martin and Turner (1986) state, "typically, these particular kinds of inquiry generate large amounts of data, which accumulate in nonstandard and unpredictable formats. The grounded theory approach offers the researcher a strategy for sifting and analysing material of this kind".

The actual procedure for carrying out the analysis by grounded theory followed the recommendations of Martin and Turner (1986). *Concept notes* were compiled from data in the interview notes which seemed to share a pattern or something else in common at a higher level of abstraction. These concept notes were added to or subtracted from as the study went on. *Theoretical memos* were written as possible explanations for the patterns observed in the concept notes occurred to the researcher, and these were also revised as the analysis went on. Example extracts from a concept note and a related theoretical memo are provided in appendix 2.

3.2 PATTERNS WHICH EMERGED FROM THE DATA

The analysis of the data using grounded theory established the existence of patterns in the design descriptions of the case study buildings. Furthermore, patterns also emerged from the interview data which enabled the development of a partial model which describes the underlying design development process which led to these design descriptions. In terms of research validity, it was found that the recollections of the project engineers were supported by the documentation in all major respects. It could be argued that this accuracy might be attributed to the engineers having read over the project files prior to the interview, in which case their recollections may have been 'contaminated'. However the engineers had no way of knowing what sort of questions would be put to them, nor even the aspects of the case which the research would focus upon. Hence it seems implausible that mere priming of this kind could explain the extent of the agreement. It is more likely that the agreement is due to the fact that the questions put to the engineers were more concerned with facts than interpretations (ie had good reliability) and that the projects were either current or very recent, and so the events were still fresh in their minds. The agreement was therefore accepted as corroboration.

This section reports on the patterns found and explains the partial model. Such is the creative nature of inductive research that it is not possible to relate (in any informative way) the actual reasoning processes by which the patterns were uncovered and the model constructed (Glaser and Strauss, 1967). However, *post hoc* justifications for these findings are provided using examples from the data.

3.2.1 Identification of the 'Feasibility Stage' Design

One of the first points to emerge from the study was that the variation in level of detail within the designs was so great that the consulting engineers found it difficult to identify the feasibility stage in the development of a project. When asked to flag the feasibility stage in the project file, a good deal of deliberation was required on the engineers' part, often involving consultations with colleagues. The factor which the engineers initially used as an indication of the feasibility stage was in fact not related to the design description at all, but was the way in which their fees were being paid. Under the traditional form of contract, payment for work done at the first two stages of the RIBA plan of work (termed 'inception' and 'feasibility') is made on a negotiated fixed fee or man-hour rate basis. Payment for design work beyond the feasibility stage is normally based on a percentage of the estimated engineering value of the job.

As the study continued, however, other design related indicators emerged which pointed to the stage of design development, but which had not been obvious to the engineers prior to the study. It unfolded that in most of the cases, the design for all areas of the building was *at least* as well developed as was indicated by the RIBA plan of work. For example, in the office projects, although there was little information on the HVAC design of the ancillary and canteen areas compared with the office areas, it was at least known what level of servicing would be provided and how much space would be taken up. In the other projects, it was also observed that there was a certain minimum level of information on those areas which had not already been considered in more detail. The engineers believed that this level of information did ensure that "the form in which the project is to proceed" was clear, which is the stated aim of the feasibility stage.

Hence, it emerged that in practice the RIBA plan of work was being used as a guide to the *minimum* amount of information which should be available at the feasibility stage.

This finding cannot, however, entirely solve the problem of placing the stage of development using the design description alone, because the plan of work is itself open to a degree of interpretation.

3.2.2 Observed Variation in Level of Detail in Design Descriptions

As is evident from the narratives in appendix 1 and the summarised information in table 3.1, there was a considerable variation in the level of detail in the design descriptions for all of the projects. However, there were significant patterns in these variations.

For the office buildings in the study (projects 1 to 4 and part of 5) the office areas were always located on the plans, although there may not yet have been any distinction made between open plan space and partitioned offices. Layouts for partitioned spaces were never available. Approximate information on the structural grid and the dimensions of structural members was always available for the office spaces but not necessarily for other parts of the building. Areas had been set aside for service cores (for stairs, lifts, toilets and secondary plant rooms), but their precise location was not always known and there was never any detail on how these spaces would be laid out. Likewise, areas had been set aside for ancillary spaces such as canteens and their associated services, but no further layout information had been worked up.

Detailed sections on a 1:20 or 1:50 scale had been sketched on all but one of the office buildings (project 3), which as an exception was explained. These sections showed the arrangement of the HVAC systems in the open plan/modular office spaces, and carried information such as approximate duct sizes. There were no such sections for ancillary areas or service cores.

In spite of the considerable variation in the level of detail, it could therefore be said that there is a 'typical' form of design description for office buildings at the feasibility stage.

The other projects and the factory area of project 5 showed a similar variation in the level of design detail. On the plans, structural information was available in some areas, and the layout of some spaces had been determined while others had merely been given an aggregate gross floor area allocation. Detailed sections had been sketched for some areas, while there was no similarly detailed information on other parts of the building.

If the levels of design development are envisaged as forming a hierarchy where higher level problems are split into sub-problems at a lower level and so on, (there is general agreement among design theorists that design hierarchies exist as *technical notions*) then the findings of the study to this point are that the various *branches* of the hierarchy are not developed uniformly through time. The RIBA plan of work does, however, determine the *minimum* level of development in the hierarchy at the feasibility stage. This is illustrated in figure 3.1.

These findings are at considerable odds with what might be expected according to the RIBA plan of work, and bore further investigation.

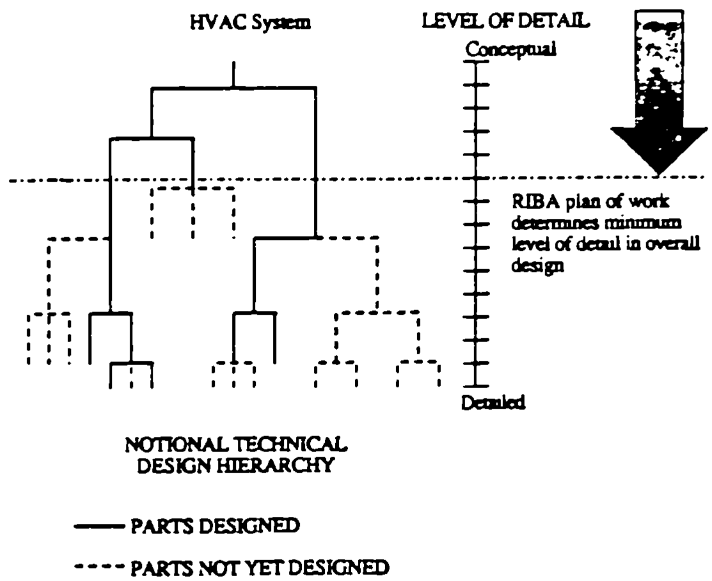


Figure 3.1 A pictorial representation of non-uniform design development

3.2.3 The Process Causing the Fragmentary Development of Detail

Explanations for the variations in level of detail observed in the 'snap-shot' of the feasibility stage design description were constructed by going back into the project files and by questioning the consulting engineers. Consistent patterns emerged across the eight case studies which allowed certain inferences to be drawn as to the design process which led to the fragmentary development of detail in the HVAC designs. Here, five general statements are presented which summarise each set of inferences as a theme and form a description of this process:

1. *The Engineers Break Problems Down into Sub-problems*

2. *The Engineers Identify Critical Details of the Design*

3. *The Engineers Investigate the Effects of Critical Details*

4. *The Engineers Judge whether the Critical Details are Sufficiently Understood*

5. *The Engineers Execute this Process Iteratively*

1. *The Engineers Break Problems down into Sub-problems*

It is clear from the evidence of the case studies that the HVAC engineers broke down the task of designing the services for the buildings by first considering the different functional areas to be accommodated. In the office projects, there was consistent separation of the HVAC systems for private offices and conference rooms, open plan/modular areas, circulation areas and canteens. In the other projects there was a similar concern with treating different areas as substantially different problems, although the engineers were always aware that ultimately the HVAC designs for the different areas must be compatible (eg to utilise central plant efficiently).

2. *The Engineers Identify Critical Details of the Design*

The step in the design process which followed the breakdown of the functional areas was the identification of those areas which involved critical design details. In the office building projects investigated, the design of the HVAC system for the open plan areas was always identified as critical, but the design of the HVAC system for the canteen was not. One engineer's comment was typical: "We know that we'll put in a constant volume packaged system in the canteen, and maybe some perimeter heating - finned tube type things - at the windows. There's definitely enough room for this and we've left enough plant space adjacent as well so the canteen is no problem".

On other occasions, the distinction was made on the basis of a critical case - that is the logical argument 'if it works here then it will work everywhere else'. In project

7 the design of the animal keeping rooms had been critical because of the large air ducts involved, however the design of the services for the near-by dissection laboratories was not critical. The engineer said "We knew we'd need to drop the corridor ceiling height right down to fit the ducts in for the animal rooms. Even at that it was a bit of a push. Anyway, because we'd done that there was no question about there being enough space for the ducts going to the dissection labs...we just knew it would be okay."

3. The Engineers Investigate the Effects of Critical Details

Investigating the effects of critical details often entailed a degree of design work. For example, in some office buildings, to know if air conditioning as against mechanical ventilation was required, the engineer had to perform heat gain calculations. To then know if sufficient space had been allowed for the required HVAC services, he had to know the generic system type(s) which might be used and so on. He was driven to perform a degree of design work, not for its own sake, but because it was requisite to his investigations.

In particular, at the beginning of the feasibility stage, it was normal for the engineers to deliberate over a number of possible generic HVAC design solutions for what they believed to be critical areas of the building as envisioned by the architect. The exceptions to this were when the engineer was constrained by the client as in project 3, or when the HVAC application was so specialised that it was believed to automatically imply a generic system type, which was the case in project 7. It should be noted that, in general, the engineers had not been involved in any serious consideration of alternative architectural designs for the building in the first instance. It was generally accepted by the engineers that, on typical projects, HVAC design considerations do not dictate the architectural design, but merely affect it in certain ways - chiefly in connection with space allowances. The exception was project 1, which was undertaken by a multi-disciplinary firm, and where there had been genuine consultation on the impact of building concept on the HVAC requirements.

In critical areas this deliberation process often culminated in a pro and con type analysis, with no attendant design work as such, to narrow down the range of options. An example of such an analysis for the open plan areas of an office building is given in figure 3.2. This analysis was conducted by the engineer acting, as he saw it, in the client's interest, by attempting to identify the HVAC system which was 'best' in the circumstances. Sometimes this initial selection process would result in the emergence of a clearly preferred system, but at other times two systems would pass this stage and require a more detailed analysis to facilitate the choice between them. In either case, it was then necessary to go on to test the technical feasibility of the system(s) by ensuring that the HVAC plant for the areas which were believed to be critical could be accommodated within the building, and executing calculations or simulations to test out the expected performance of the systems.

SYSTEM	ADVANTAGES	DISADVANTAGES
VAV with perimeter re-heat	High fresh air volume Low noise output Good space flexibility Low maintenance requirement	Very high space requirement High capital cost
Fan assisted VAV	Low running costs Moderate fresh air volumes	High space requirement High capital cost Moderately high maintenance requirement
Fan coil with Mechanical Ventilation	Good space flexibility Low capital cost Low space requirement	High noise output High maintenance requirement Low fresh air volume
VRV with Mechanical Ventilation	Very low space requirement Low capital cost	High noise output High maintenance requirement Low fresh air volume

Figure 3.2 Example of a pro and con analysis of different air conditioning systems

4. The Engineers Judge whether Critical Details are Sufficiently Understood

Where the effects of certain aspects of the design were judged to be well enough understood, and so appropriate allowances for space, location etc could be made in the overall feasibility stage design, the designers did not believe it necessary to go into any further design detail. For example, in all of the projects studied, the space required for plant rooms was calculated from rules of thumb. The engineers were sufficiently confident in this information not to feel compelled to start designing the layout of the plant rooms in detail.

In project 5 the client had a comprehensive briefing and design guide which contained, among other things, some detailed specifications for the HVAC provision in the offices and factory production areas. There is no doubt that the servicing of the production facilities was a critical aspect of the design, but following his investigations of these critical details, the engineer felt that the client's guide was so specific in that regard that no immediate design work was required: "Everything was in there down to the size of the pipe connections they needed for their machines. We knew we could distribute the services to meet their requirements in the production spaces. The only thing we had to check on was how we'd tie it all together in the main runs from the plant rooms."

The proviso by the engineer in the above quotation is illustrative of a critical detail within a critical area which was not sufficiently understood. The engineers tackled these details by further breaking the design problem down into sub-problems (sub-sub-problems as it were) and then once more going through the process of identifying critical details (sub-details), investigating their effects, and so on. In the cases studied here, these were details such as the distribution of major duct work or the means of accommodating items of plant (eg fan coil units). Such details were usually considered important unknowns which had to be investigated because the feasibility of the wider HVAC design depended upon them, and they had an influence on the allowances made in the overall design of the building (eg floor to ceiling height, external solar shading). It was these critical and unknown details which were the subject of the larger scale sections through the building and the calculations and simulations performed by the engineers.

5. *The Engineers Execute this Process Iteratively*

As can be seen above, the process of identifying critical details, investigating their effects and so on has begun to execute a second cycle, but on a finer level of detail. In the course of carrying out these more detailed design tasks, however, the engineers could come across yet another set of sub-details which were critical to the larger design problem and so require investigation. For example, in project 4, it was proposed that the supply of air to the open plan office-spaces would be upward via grilles in the raised floor. The engineer was concerned that there would be sufficient mixing of the air within the space, however, and that annoying draughts could be avoided. This then forced him to investigate various forms of grille type and the arrangement of these in the floor. In project 8 the engineer would have preferred to locate the chiller plant on the roof of the library building for convenience in terms of pipe runs and other practical matters, but he was concerned about the possible noise problems this could cause. Resolving this question had required obtaining detailed information on the typical sound and vibration levels produced by chillers of the kind he anticipated using, together with details of the building structure anticipated by the architect and structural engineer, and then approaching a specialist acoustics consultant to perform calculations which would predict the likely noise levels in the sensitive areas of the library.

This iterative process could, in principle, continue ad infinitum, however in practice as a design problem is broken down into smaller and smaller parts, the solution of the smaller design problems becomes progressively more trivial (although the wider scale coordination of these in the context of the overall design becomes more difficult at the same time). For this reason, there would seem to be a manageable number of iterations before a level is reached at which the designer is satisfied that the effects of that critical detail on the preceding level of detail is sufficiently well understood. In the cases studied here, it was estimated that by the feasibility stage the engineers had gone through this iterative process between three and four times at most before stopping. This estimate is very tentative, however, as the distinction between one level of detail and another is arbitrary.

3.2.4 Factors Which Influenced the Development of the Process

As was to be expected, the way in which the engineers actually carried out the above steps of breaking the design problem down into sub-problems, identifying critical details and so on was highly complex, and influenced by a myriad of factors. Understanding these methods and influences in depth goes to the heart of the designer's skill (as referred to by Simon and discussed in section 2.2.2), and is the 'holy grail' of descriptive design models. It was not expected, therefore, that a preliminary study such as this using high level data obtained from project files and unstructured interviews would lead to a comprehensive understanding. However, this research was specifically concerned with the evolution of HVAC designs at an early stage, and so within this restricted area it was nevertheless possible to obtain an insight into the more obvious factors which influenced the engineers during the process.

Client Requirements

In the projects studied here, it was observed that the way in which the client requirements were presented and the degree of detail accompanying them could have a significant influence on the way in which the engineers identified and dealt with critical details.

In some instances, the client stipulated certain requirements which the design had to address if it was to be accepted. These requirements could be functional, such as the stringent stipulations on temperature, noise etc for the animal keeping rooms made by the client in project 7 and the detailed specification of the process equipment for the production spaces in project 5, or they could be of the "bee in the bonnet" variety as it was termed by one engineer, such as the audio-visual room in project 1.

In project 7, the strict requirements set down immediately identified the design of the HVAC systems for the animal keeping rooms as critical. Moreover, the engineer was not satisfied that enough was known about how to meet these requirements to make allowances in the general design. He therefore proceeded to develop certain aspects of the design of the animal rooms in detail.

The engineer on project 5 identified the design of the HVAC services in the production spaces as critical because the client was so specific about their requirements. However, the amount and type of detail supplied by the client was such that the engineer felt that the effects of these critical areas was sufficiently well understood for that stage in the general design and so no detailed design was necessary.

In project 1, the detail in the design of the audio-visual room was at a stage far in advance of any other part of the building. The engineer stated that he believed the design development to be more in keeping with stage E (detailed design) than stage B of the RIBA plan of work: "if they [the client] accept the design, and we don't come across any problems ourselves, then I don't think we'll need to do much more on this". This was in spite of the fact that the area was not considered at all important from a design detail point of view. The engineer explained that the client was particularly enthusiastic about having the audio-visual suite as part of the new headquarters, and therefore an important "selling point" of the entire design was to address this area as fully as possible: "This has got more to do with knowing which side our bread is buttered on than designing the services for the building".

Externally Imposed Restrictions

Just as client requirements obviously influence design, so too do factors such as the shape and condition of the site and planning regulations. Restrictions of this kind imposed on the engineers from outside were another influence on the way in which they decided whether or not to proceed to develop a more detailed design description. In project 3, for example, the engineer was concerned about the planning restrictions relating to roof mounted plant, and so some detailed discussions with the architect had taken place. In project 7, the design of the fume cupboard exhaust flues had to meet strict regulations to ensure that the gasses would be discharged into the atmosphere without causing a hazard to those in nearby buildings. To do this, the engineer had been forced to decide at a very early stage on the location of the flues on the roof of the building and had then commissioned wind tunnel tests using a scale model of the site.

Sufficient Information for Detailed Design

The requirement to consider detailed aspects of HVAC design at an early stage led to problems on occasion, because sometimes the detailed design required commensurately detailed information from other designers. For example in projects 1 and 4 computer thermal simulations were carried out to test the proposed HVAC system design in the open plan office areas. This simulation required detailed information on the construction of the outside wall such as the amount and type of glazing, the thermal transmittance of opaque panels and so on. However the architects were frequently not in a position to provide that information at such an early stage in the design. Another example of such a problem which applied to all of the cases studied was the design of the structural frame. Once again details such as the likely dimensions of beams and columns (which restrict the space for services) were required before the structural engineer was in a position to produce detailed calculations.

In these instances, rules of thumb or building regulations were used as 'default' values, but with figures which, as one engineer put it, "erred on the safe side so that when they do come up with the details later on there are no nasty surprises". All of the engineers acknowledged that the consequence of this was to over-design, however they could see no alternative to this approach, and indeed considered it as a form of intuitive risk analysis: "You might end up over-designing - in fact you often do - but you try to build enough flexibility into your system so that you can tweak it down later on. You have to weigh up the cost of over-designing a bit against the cost of making some kind of disastrous mistake in underestimating. It can cost a lot to fix those!".

3.3 CONCLUSIONS OF THE PRELIMINARY STUDY

This preliminary study has revealed much about the approach of engineers to the design of HVAC systems at the feasibility stage. The findings are quite general and provide valuable insights into the design process which informed the development of the larger research project in ways which shall become clearer in later chapters. These findings are summarised below in section 3.3.1. Among these findings, however, are results of more currently obvious relevance to the specific research aims of the thesis, and these are highlighted in section 3.3.2.

3.3.1 General Conclusions

It was found that the design descriptions studied here exhibited a variation in level of design detail. This is contrary to the implications of the RIBA plan of work and the BSRIA guide for engineering services design, but not unexpected given the theoretical background provided by the design research literature discussed in chapter 2.

The extent of the variation found in level of detail in the designs was considerable, ranging from the nature of the information required by the plan of work at the feasibility stage (choice of generic HVAC system type) as a minimum, to highly detailed sketches of sections through certain parts of the building and analyses of computer thermal performance simulations. Despite this variation in level of detail, patterns found in the variations indicate that there is a 'typical' form of design description for office buildings at the feasibility stage. Moreover, similar consistencies in the designs for the other building types considered in the study suggest that the technical design process behind this is not entirely idiosyncratic, but points to a strategic approach to the use of detail which is shared by the engineers, and which is more directional than the 'whirling' type of activity suggested by Lawson and Hickling (section 2.1.3).

By going back into the project files and questioning the project engineers, certain inferences were drawn as to the process which leads to the fragmentary development of detail in HVAC designs. Central to this process is the notion of the *critical detail.* In each design, there were certain details which were considered critical if the design was to be technically feasible and/or acceptable to the client. These details might be obvious from the brief and discussions with the client or architect, or considerable experience might be required to recognise them. However they are perceived, it is the need to investigate the effects of these critical details so that appropriate allowances can be made in the overall feasibility stage building design which leads to the apparently premature plunge into detailed HVAC design in selected areas. The process which the engineers go through was described as follows:

1. *The Engineers Break Problems Down into Sub-problems*
2. *The Engineers Identify Critical Details of the Design*
3. *The Engineers Investigate the Effects of Critical Details*
4. *The Engineers Judge whether the Critical Details are Sufficiently Understood*
5. *The Engineers Execute this Process Iteratively*

The design activity involved within stage 3 of this cyclical strategy seems to begin with a two stage appraisal process. First an assessment is made of the most appropriate fundamental HVAC system provision for each area of the building from a broad generic range (heating and natural ventilation through mechanical ventilation to air conditioning), then secondly an assessment is made of generic system types from within this provision category, provided that the engineer does not believe that the preferred choice of system type is obvious from the outset. Once one or two candidate systems have been chosen, further design activity might follow in subsequent iterations of the cyclical strategy if the engineer requires more information on certain critical details such as major duct runs and accommodation for certain items of plant.

Hence it seems from this study that a hierarchical process of HVAC design proceeding from the conceptual to the specific (similar to that assumed by the RIBA and BSRIA guides) does exist, however it is not uniformly hierarchical in that at any point in time it may be found that some details have been considered while others have not. The attention of the designer is directed through the technical design hierarchy by the strategy above, the principle motivation at the early stages of design being *investigative*, rather than to complete the design for its own sake. In terms of design completion, it was found that the RIBA plan of work influences the strategy by acting as a guide to the minimum amount of information which should be available at the conclusion of the feasibility stage.

This technical design process is relatively self-contained, where the HVAC designer's interaction with other members of the design team is minimal. This sometimes involves making assumptions about aspects of the architectural design which have not yet been decided, indicating that the engineers temporarily treat ill-structured problems as well-structured to facilitate solution.

The foregoing suggests a rudimentary descriptive model of the technical HVAC design process at the early stages, and this model is summarized below in figure 3.3. The model is a partial one and seeks only to describe the *mode* of strategic behaviour observed as engineers move through the levels of detail associated with the vertical dimension of the two dimensional design model discussed in section 2.1.1. It does not explain the detailed cognitive process of how engineers actually recognise critical details and navigate their way through the technical design hierarchy, although the study did highlight some of the more obvious queues influencing the designer, such as the statement of client requirements, restrictions imposed by regulations and so on.

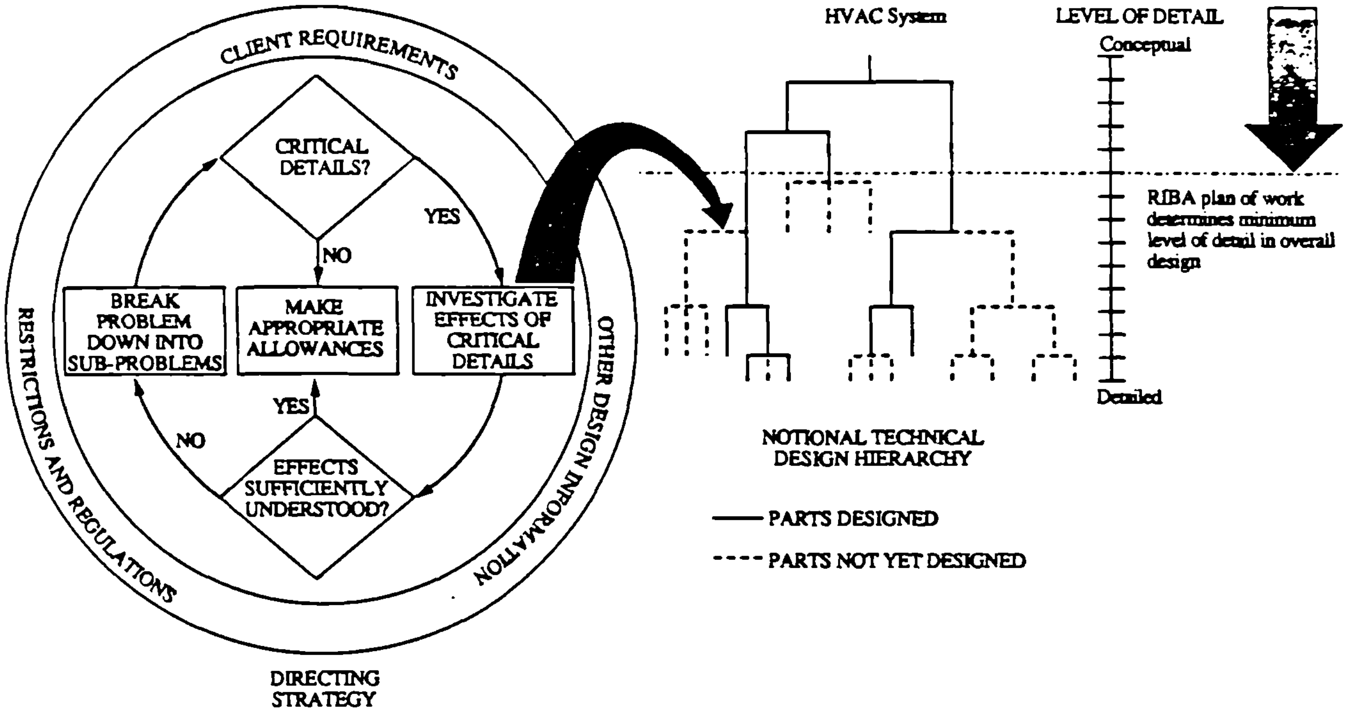


Figure 3.3 A descriptive model of the mode of strategic design behaviour leading to the fragmentary development of detail in HVAC designs

3.3.2 Specific Conclusions

In terms of the specific research aims of the thesis, this preliminary study has shown that while the development of HVAC designs appears on the surface to be very different to the top-down approach assumed by the systematic model, engineers do adopt a hierarchical process of HVAC design proceeding from the conceptual to the specific. This process of design seems to begin with a two stage appraisal in which the fundamental HVAC provision for each area of the building is assessed and then an evaluation is made of generic system types from within this provision category. This process is most intensive in the areas of the building considered to be critical, often culminating in a pro and con type analysis in which the advantages and disadvantages of system types are weighed against each other. This suggests that in the initial stages the 'typical' HVAC design process *can* be characterised as a decision between alternative generic system types. Further investigation into the scope for the use of formal appraisal methods such as weighted scoring would therefore seem justified in this context.

Furthermore, it has also been shown that 'office buildings', or indeed any other building, is not a homogeneous concept in the design of HVAC systems, but rather they have different functional areas (eg open plan/modular spaces, canteens, ancillary spaces, circulation spaces) which are treated as substantially separate problems of varying importance. In the case of office buildings, it is the open plan/modular office space which is considered the most critical area of HVAC design. This indicates that further research into the decision making process would most realistically be concentrated on this type of area.

Before proceeding to study this decision making process in greater depth however, it is first necessary to understand the theoretical background to the problem solving/decision making dimension of the systematic design model in general, and weighted scoring methods in particular.

Chapter 4

The Problem Solving/Decision Making Dimension of the Design Process

As was described in section 2.1.1, the systematic design model is two dimensional. The vertical dimension corresponds to the development of the design from conceptual to detailed design, while the horizontal dimension represents the problem solving/decision making activity which occurs at each stage of development. This chapter now looks in detail at the horizontal dimension.

Appraisal methods such as weighted scoring are placed in context within the horizontal dimension as prescriptive decision making methods. It is argued that while decision making *per se* may have a limited role in solving some design problems, it has an important place in the typical HVAC design process and so is worthy of study in this sense at least, which lays the philosophical basis for the continuing study. The difficulty in assessing the effectiveness of prescriptive decision making methods, which has implications for research methodology, is then addressed. Finally the chapter closes with a discussion on the rationale behind appraisal methods such as Cost-Benefit Analysis and weighted scoring (as normally described in the design methods literature), and provides a brief critique of those methods.

4.1 THE UNIVERSALITY OF THE DESIGN PROCESS

The optimising model of analysis-synthesis-appraisal leading to a decision is not specific to the field of design theory. The model has its origin in systems analysis, but it is also found in structured approaches to problem solving/decision making in the so-called 'soft' (ie pertaining to ill-structured problems) applications of management science, operations research and decision analysis (see for example Oberstone, 1990; Andriole, 1989; Van Gundy, 1988; Sage, 1986; Nadler, 1981; Huber, 1980; Ackoff, 1978; Newall and Simon, 1972). Moreover, some of the less context specific design methods mentioned in section 2.2.4 such as objectives trees, brainstorming and - most significantly for this work - weighted scoring are also found as methods in these other fields of study. This is not mere coincidence, but reflects the underlying similarities in the fields.

Margolin (1989) and Buchanan (1990) argue that there is no real distinction between 'design' as taken to mean the conception and planning of new products or buildings and the conception and planning of a variety of other things:

> "The central theme of design is the conception and planning of the artificial. Design provides the thought which guides the making of all products, whether by individual craftsmanship or mass production techniques: (1) material objects, (2) verbal and visual communications, (3) organised activities and services and (4) complex systems of environments for living, playing, working and learning." (Buchanan, 1990).

On the same theme, Willem (1991) asserts that design as a process is basically the same in every field of endeavour where people conceive and plan new things, the differences lie in the medium and the circumstances in which the process is carried out. From this, it may be concluded that many of the observations and discussions originating from one field are applicable in the others.

4.1.1 The Analysis-Synthesis-Appraisal Process

As discussed in chapter 2, the systematic design model is prescriptive, and seeks to organise the behaviour of designers in such a way that they become more effective. Such is the complexity of real design tasks however, that the model cannot be taken literally, and requires 'sensible' interpretation.

The literature from the various design related fields mentioned above reflects a broad consensus on a sensible interpretation of the optimising analysis-synthesis-appraisal process which takes into account the *bounded rationality* of problem solvers (Simon, 1957). This consensus model is described below:

Analysis

It is understood that the analysis stage cannot be a complete, once-and-for-all activity carried out only at the start of the process. Some initial analysis of the problem to be solved will be carried out until the problem solver believes he has acquired sufficient understanding about the problem that he may begin to synthesise solutions (Howard, 1988). It is implicit in this that initial, tentative design objectives and criteria by which attainment of these objectives can be assessed will have been evoked in the mind of the problem solver.

Synthesis

The notion that an optimal solution can somehow be synthesised from the systematic search of a solution space directed by a complete set of optimising criteria was not promulgated in the context of ill-structured problems after the 1960s (Churchman et al, 1975; Quade, 1975). Instead, the synthesis stage is generally treated as one in which a range of possible solutions to the problem, as it is so far understood, are generated.

It is recognised that this stage is probably the most important in the entire process, but it is also believed to be the stage least accessible to structured assistance. This is because the generation of solution ideas is basically a creative activity, and as such is inspirational, rather than a purely cognitive and disciplined process. This view is reflected in the methods prescribed to encourage creativity such as brainstorming and synectics. Such methods are intended to break the mind-set of the problem solver by (somewhat paradoxically for a systematic approach to problem solving) introducing a degree of irrationality into the process, so encouraging the generation of as wide a range of solution options as possible.

Appraisal

It is the appraisal stage which receives the most attention in the literature because, unlike creativity, appraisal is seen as a disciplined and rational activity and therefore open to the application of structured methods. The optimising model has as its goal the selection of an option which is in some sense the best *in the perception of the decision actor*. This requires estimating the value of every alternative in terms of the expected performance on the range of relevant (and often conflicting) criteria, and selecting the one which has the highest aggregate value. This is the function of methods such as Cost-Benefit Analysis in economics and, of course, weighted scoring.

The Iterative Nature of the Process

The systematic model recognises that problem solving/decision making is an iterative process. However, the full nature of this iteration is not acknowledged across all the fields of study. It is usually assumed that feedback occurs primarily between the appraisal and synthesis stages, as the favourable or unfavourable assessment of certain options informs the search for more solutions. The problem definition, however, remains unchanged. Empirical studies in the field of design theory (see section 2.1.2) and works in the field of management problem solving/decision making (Koestler, 1978) indicate that an even more fundamental feedback process occurs between the synthesis stage and the analysis stage, as the generation of possible solutions can influence the definition of the problem. This is seen as a crucial element in successful creative problem solving by some writers (Van Gundy, 1988) and has been accorded a prescriptive status so far as such creative behaviour can be 'prescribed'.

4.1.2 'Problem Solving' and 'Decision Making'

Thus far, the inference has been that 'problem solving' and 'decision making' are synonymous, following the use of the terms by Markus and Maver (see section 2.1.1). In the literature from the various fields reviewed, many writers treat the terms as virtual synonyms, while others do not. Unfortunately, the definition adopted is almost always implicit rather than explicit, and sometimes seems poorly understood by the writers themselves. In this thesis, *problem solving* will be regarded as the entire dynamic process of analysis-synthesis-appraisal, ie design is viewed as a problem solving activity. Following the dominant use of the term in the field of decision analysis, *decision making* will be subsumed within problem solving as the rational process of setting performance criteria and evaluating previously generated design options against these criteria.

While recognising that it is only part of the wider (and richer) process of design, this thesis is primarily concerned with decision making as defined above.

4.1.3 Is the study of decision making in design a trivial pursuit?

It has been argued by some writers that once a problem solving situation has been reduced to a decision making situation (to use the above terminology) the major difficulties have already been overcome (Quade, 1975; Checkland 1981; Von Winterfeldt and Edwards, 1986; Baron, 1988). The initially ill-structured problem has already been identified and divided into well-structured sub-problems, the sub-problems have been clarified to an extent through iteration, and a range of solution options for the sub-problems has been generated. All that remains is to choose between the options.

The degree to which this stage of the design process may be considered trivial in comparison to the others depends upon the nature of the problem solving situation. Sometimes, although there may be a vague perception that 'something is wrong' or 'something needs to be done', the problem cannot actually be identified. Von Winterfeldt and Edwards (1986 p 31) refer to such problems as an "unscratched itch". On other occasions, the problem perception may be clearer, but no solution options can be generated. In such cases, then it is true that the selection of one option from a range of options seems a distant and relatively unimportant stage in the problem solving process.

However the problems to which these writers refer are generally novel problems. The design of HVAC systems, which is the specific area of interest in this thesis, may been seen in that light if a highly novel design is required for a highly novel building, but in practice it would seem that such instances are rare. According to the BSRIA design guide (BSRIA, 1990) and implications in other HVAC design literature (eg CIBSE, 1986; ASHRAE, 1991; Legg, 1991; Stanford, 1988) more typically, the early stage design of HVAC systems can be characterised as a decision between alternative generic system types, and this was confirmed by the preliminary study reported in chapter 3. In this context at least then, a study of this decision task and ways of facilitating it such as by using the weighted scoring technique seems a reasonable and worthwhile undertaking.

It may be argued, however, that adopting such a position makes too much of a concession to how HVAC designers actually work. The optimising model is intended to encourage designers to consider a wider range of options thereby ameliorating the danger of perpetuating inferior designs - yet the decision making situation has been characterised as the selection of an HVAC design from a *known* range of generic system types.

Like so many other issues in the study of design, this is ultimately a matter of degree and there is no definitive answer to this criticism. It is noted that an increasing number of others in the design research community are adopting this position in discussing the role of decision aids in design (Blanford, 1993; Beheshti, 1993; Dobias, 1990). This fact does not, however, constitute an explanation. More substantially, the position adopted here may be justified in two ways:

Piecemeal Engineering

Faludi (1986) applies Karl Popper's views on 'piecemeal engineering' (Popper, 1972) to planning and design. He points out that designers should not seek to innovate in large jumps, but should proceed more prudently in incremental steps to avoid making large errors with untried technologies. Within this constraint, however, he maintains that designers should attempt an optimising approach to problem solving. The typical practice of HVAC designers in confining the selection of HVAC system to one of a range of known generic system types can be interpreted in this manner, with incremental progress occurring as new systems appear on the market and become 'known types'.

Methodological Generality

It is anticipated that the selection of a system from a range of known types is a useful starting point in investigating the suitability of an appraisal method to the task of HVAC system selection in general, the principles of which may also be applied to novel designs.

4.2 MODELS OF THE DECISION MAKING PROCESS

Simon (1957) puts forward two theories about the fundamental strategies which may be adopted when making decisions, and these have endured as the dominant paradigms of decision making (Klein and Calderwood, 1991). He terms these two fundamentally different strategies of decision making *maximising* and *satisficing*. The maximising strategy has as its goal the selection of an option which is in some sense the best, and is generally associated with an overall optimising approach to problem solving. As previously described, this requires estimating the value of every alternative in terms of the expected performance on a wide range of criteria. The satisficing strategy on the other hand has as its more limited goal the selection of an option which meets a minimal set of requirements and is therefore 'good enough'. This only requires assessing a number of alternatives in any random sequence until one which fulfils the minimal requirements is found.

4.2.1 Heuristic Decision Making Strategies

Subsequent to Simon's early work, several writers on decision making have pointed out that although they believe the fundamental paradigms are correct, there is no straightforward distinction between maximising and satisficing (Janis and Mann, 1977). Rather, these strategies represent extremes of a large variety of possible decision making strategies. A number of writers have observed numerous intermediate decision making strategies which are based on heuristic elimination, and interest in these strategies has increased in recent years as it was realised that 'expert judgement' is substantially heuristc in nature. Understanding these heuristic reasoning processes has enabled researchers to take the first steps towards imitating expert reasoning processes in expert system computer programmes (Liebowitz and De Salvo, 1989).

Montgomery and Svenson (1976), Olshavsky (1979) and Payne *et al* (1988) report a number of heuristic elimination processes which have been observed in studies of decisions with several criteria. The number of individual strategies is large, however three of the more fundamental strategies which form the basis for other, more complex variations can be decerned.

The Conjunctive Strategy

One of the simplest strategies is the *conjunctive* strategy (Einhorn, 1971) in which minimally acceptable standards are set for each criterion and all options which fail to exceed any of these standards are rejected. If more than one option passes this assessment, the evaluation process is repeated with raised standards until only a single option remains.

The Lexicographic Strategy

The *lexicographic* strategy (Fishburn, 1974; Hogarth, 1980) prescribes the selection of the option which performs best on the most important criterion. In the event of a tie the comparison is repeated based on the second most important criterion and so on until only one alternative remains.

The Elimination by Aspects Strategy

A slightly more complex strategy is what Tversky (1972) calls the *elimination by aspects* approach, which has also been termed the *split and prune* strategy (Pearl, 1984). Using elimination by aspects, decision making is a sequential narrowing-down process. Starting with the most important requirement, all salient options that do not satisfy the requirement are eliminated. The process then continues for each requirement in turn in descending order of importance until a single option remains.

These processes are more comprehensive than satisficing as they involve the consideration of an entire range of options and attempt to find one which is 'best', however the method by which they do this is much less rigorous than maximising. The lexicographic approach will generally fail to consider more than one or two criteria, and the reliance on performance thresholds alone in all of the strategies means a failure to account for how much the performance of an option exceeds the threshold on each criterion. These deficiencies mean that the strategies are *non-compensatory* (ie poor performance on some criteria cannot be compensated for by good performance on others) and so can lead to the ultimate selection of an option which is, in an aggregate sense, inferior to some of those which have been eliminated.

Finally, Etzioni (1967) outlined a combined strategy which he refers to as *mixed scanning*. In essence, this involves the use of simpler evaluations such as conjunctive analysis and then an elimination by aspects approach at the early stages in a decision situation to eliminate those options which have major deficiencies in important requirements, then switching to a maximising type approach when the number of options and the number of criteria to be considered have been reduced to more manageable proportions.

4.2.2 The Application of Decision Making Strategies

Simon (1957) asserts that in practice the dominant strategy is that of satisficing, because people do not have the cognitive ability, the time or the resources to maximise. This assertion has also been made by other researchers. Miller and Star (1967) point out that a person attempting to maximise will be overwhelmed by information, as the number of items of information on which the decision must be based will usually exceed seven, which is generally accepted to be the limit of man's unaided capacity to process information in immediate memory (Miller, 1956). Some evidence can also be found in other social science disciplines to indicate that, in addition to man's cognitive limitations, practical constraints usually militate against the adoption of a true maximising strategy. These are reviewed in Janis and Mann (1977).

Contradicting this view to some extent, however, are yet more works which suggest that people adopt contingent decision behaviour, where the effort expended on the decision process depends upon the importance of the decision and the effort required in decision making (Townsend, 1970; Beach and Mitchell, 1978; Payne, 1982; Gilliland *et al*, 1993). For unimportant decisions, particularly those which can easily be reversed, satisficing may be seen as adequate because there is no point in expending a great deal of effort upon them. However for important decisions, and especially those which are difficult to reverse, it may be thought that it is worth while expending considerably more effort and attempting to maximise.

It would seem that, to date, no dominant descriptive model of how people make decisions in practice has emerged (Sainfort et al, 1990; Gilleland *et al*, 1993).

4.2.3 The Prescriptive/Descriptive Dichotomy in Decision Models

Just as in design theory, there is a prescriptive/descriptive dichotomy in the theory of decision making, although the dichotomy is not quite as clearly defined. The satisficing strategy and the other simpler decision making approaches are viewed as predominantly descriptive, as although many people are observed to make decisions in these ways, such strategies would not be recommended for important decisions. On the other hand, maximising is predominantly prescriptive, because it encourages decision makers to be more vigilant information processors as Janis and Mann (1977) put it, and therefore less likely to make errors. The mixed scanning strategy of Etzioni could be regarded as both descriptive and prescriptive, as it advocates a strategy which seems to make normative sense, and at the same time has been observed in use in empirical studies.

4.3 THE EFFECTIVENESS OF PRESCRIPTIVE DECISION STRATEGIES

Once again, the existence of a prescriptive/descriptive dichotomy raises the question of how to know if a prescriptive model actually offers an improvement over unaided ways of working. It has already been indicated (see section 2.3) that judging a prescriptive model on *ad hoc* outcomes is problematic because the weak nature of such a model means that using it requires informed interpretation and success cannot be guaranteed. Moreover, 'success' is itself difficult to define in making decisions which have no objectively 'best' outcome. In the absence of an objective outcome based success criterion, decision analysts advocate a more eclectic approach to the validation of prescriptive decision aiding methods. The general principle behind this approach is that in the absence of a conclusive *outcome* based measure of effectiveness, the quality of the decision *process* must be used as the measure of effectiveness (Howard, 1988; Watson and Buede, 1987; von Winterfeldt and Edwards, 1986; Janis and Mann, 1977).

4.3.1 The Rationality Principle

In Western philosophy, it is intuitive to look upon a 'correct' decision as one which is arrived at *rationally* (Watson and Buede, 1987; Janis and Mann, 1977). Sage (1981) summarizes the various forms of rationality which have been posited from time to time and highlights that, in terms of decision making, rationality is primarily concerned with consistency. Consistency, in turn, is defined by a set of rules which it is agreed should govern the choice between options in the decision situation. It follows then that an effective prescriptive decision model should set down the rules governing consistent choice and make them sufficiently explicit such that a decision actor can satisfy his need to be rational by following the rules.

This, in essence, is the purpose of maximising decision strategies, but different maximising methods (such as Cost-Benefit Analysis and the variations of weighted scoring) are founded upon different rules. There must therefore be a question over which set of rules is best. Ultimately, there is no objective answer to this because the rules are axiomatic, and it must fall to a combination of abstract logical argument and assessment by the individual decision actor to determine which set of rules is most valid in a theoretical and normative sense.

4.3.2 Theoretical and Normative Validity

Detailed discussions on the issues relating to the validity of decision making methods have been presented by several writers (eg Gass and Torrence, 1991; Hobbs *et al* 1991; Hobbs, 1986; von Winterfeldt and Edwards, 1986; Stewart and Ely, 1984; Wierzbicki, 1983; Pitz et al, 1980). In terms of the suitability of a method to a particular task, these discussions highlight two main (but interconnected) forms of validity which shall be termed here as *theoretical* and *normative*. (Note that there is no agreed terminology within the decision analysis community on this issue, therefore although the definition of these terms has been derived from the literature, their precise meaning must be to some degree unique to the thesis).

Theoretical validity refers to the logical consistency of the decision making method, and is determined primarily by abstract logical argument. A method which is theoretically valid should not contain any internal inconsistencies in the algorithm, variable definitions and so on used in the method, and the evaluations made by the method should vary in appropriate ways as theoretically relevant problem characteristics are altered (eg the importance of certain criteria), and not vary at all if theoretically irrelevant characteristics are altered (eg the way in which questions are phrased to elicit the importance of criteria).

Normative validity is achieved if the method is a sound basis for making decisions, and is determined primarily by the decision actor. He must be satisfied that the model of the decision situation embodied in the method algorithm is appropriate for the decision making task, and consequently he should be able to explain and defend the evaluations made by the method.

These forms of validity must be considered together, because separately they are incomplete. As Wierzbicki (1983) points out, theoretical validity of the method algorithms "...while mathematically elegant and possibly useful to the decision analyst, does not mean anything in application until checked empirically...(A)nalysts who use certain mathematical concepts extensively are apt to believe that those concepts have some independent existence in the real world".

This concern is not restricted to 'soft' applications, but applies equally to traditional 'hard' methods in operations research, systems analysis and management science:

> "The reader might ask, when is a model valid? The answer is, simply, when the user is convinced of its usefulness in a given situation." (Grinold and Marshall, 1977 p251)

> "Either an executive believes that the operations research representation of the problem is valid or the executive dismisses the results as worthless." (Wagner, 1975 p587)

Nevertheless, assessment by the actor alone is not a sufficient measure of the validity of a prescriptive decision model. Actors faced with complex appraisals may be seduced by methods which are simple to use and have an apparently authoritative basis, yet may be found to be seriously deficient in theoretical terms. This has been observed in the case of computerised decision support systems with attractive graphical 'front ends' (Adelman, 1992). It has also been noted that, in situations where a decision analyst acts as a facilitator for the actor in using a particular prescriptive model, the perceived success of the analysis may owe more to the social skills of the analyst than to the validity of the actual methods used (Fischhoff, 1980; Lock, 1983).

4.3.3 The Constructive View of Prescriptive Decision Making Methods

Watson and Buede (1987) make one of the most useful attempts to deal comprehensively with the above facets of validity. They combine the above mentioned concepts of validity and rationality in suggesting that prescriptive decision making methods are *constructive*. They assert that the model of the decision task embodied within the method does not represent some objective 'reality', nor do the preferences inherent in the method algorithms and variables represent a latent preference structure within the actor's mind. Rather, they argue that the prescriptive method is a construction which satisfies certain principles of logic and consistency that can be called 'rational' and which is consistent with those judgements which the actor can express confidently.

Thus far, this position is in agreement with that of Phillips (1984) whose theory of *requisite decision models* has been highly influential in the field of decision analysis. He argues that a prescriptive model can be considered valid if it is *requisite*, by which he means it is sufficiently comprehensive for the actor to confidently make a decision based upon it. It may be suggested, however, that in making this the acid test Phillips is relying too heavily upon normative validity and underemphasising theoretical validity. Watson and Buede point out that the situation can be imagined where a 'requisite' model might embody implications that the decision actor would not accept and yet they remain undetected because by chance these implications were not tested. The corollary of this is that a test of validity would involve checking as many of the implications of the decision model as possible against what the decision actor knows about his preferences, and determining if any clashes occur.

In practice, however, the approach to this must be pragmatic. Firstly, the number of possible checks in a completely comprehensive assessment could be unfeasibly large for a model of realistic size; and secondly, the decision actor may find it difficult to make some of the judgements required in such an assessment since, after all, the reason for building the model in the first instance is that unaided holistic judgement is difficult. In part answer to the first problem, it might be argued that there will usually be a number of factors within a model which are more influential than others in determining the outcome of an evaluation, and so the validity assessment should concentrate on these. In response to the second problem Watson and Buede (p274) comment:

> "We do not have to agonise over whether we *really* prefer option A to option B; if we feel unable to articulate a preference between them, and yet, as a consequence of our other judgements and the principles of rationality that we wish to adopt, the model tells us that A should be marginally preferred to B, then we will be happy with that synthesised judgement. In this regard, the model is valid, not because we know that we prefer A to B, but because there is no articulable preference between the options that clashes with the implications of our model."

That is to say, the actor asks himself (or is asked by an analyst) whether certain implications of the model lead to contradictions with other intuitive judgements which he is confident in making.

Further discussion of validity issues in the context of specific assessment methods is presented in chapter 7. It will be appreciated from the foregoing, however, that the issue of validity is far from fully worked out, and so the researcher in decision analysis faced with making an assessment of validity must accept imperfections in the research methodology (von Winterfeldt and Edwards, 1986).

4.4 MAXIMISING WITH SEVERAL CRITERIA AND SATISFYING THE RATIONALITY PRINCIPLE

Making a choice between a number of options and being satisfied that the decision is rational is straightforward if there is only one criterion governing the selection. For example, if a decision actor must decide upon which motorcar to buy from among a range of options, it may be that selection of the car with the lowest price is the only decision criterion. However few real problems are as simple as this. More usually a number of criteria which conflict are involved in the decision. To extend the previous example, in addition to differing in price, the cars under consideration will almost certainly also differ in terms of power, comfort, space and so on. If the actor wishes to identify the *best overall* car, then he must adopt a maximising decision making method where the value of these other criteria can also be taken into consideration in some consistent and compensatory fashion. This form of performance versus cost decision is familiar to everyone. Moreover, these decisions are often complex, involving more criteria than an individual can consider unaided.

To assist actors to think rationally when faced with such decisions, it was posited in section that a set of rules governing consistent choice is required. This section discusses the minimal axioms of consistent choice which *any* maximising decision method should satisfy. It is then explained that maximising decision methods, which must be compensatory, must address the problem of expressing incommensurable and intangible criteria in a common measuring denominator. Cost-Benefit Analysis is probably the best known technique for dealing with this problem, and so it is discussed before proceeding to describe in more detail the simple weighted scoring method introduced in chapter 1.

4.4.1 Three Axioms of Consistent Choice

Three fundamental axioms of consistent choice which constitute the minimal requirements for rationality are set down here. These statements are colloquial versions of the more formally stated axioms found in the decision analysis literature, (see for example Raiffa, 1968; Brown et al, 1974; Hooker et al, 1978) but they are sufficiently rigorous as an introduction to the concepts involved. Although necessary, the axioms are not, however, sufficient for defining rationality, and it will be shown in chapter 5 that mainstream decision analysis makes some additional requirements.

Transitivity of Preference

If there are three options A, B and C, and A is preferred to B and B is preferred to C, then A must be preferred to C. Furthermore, if the rule applies for three options, then it applies for any greater number.

Transitivity of Indifference

If there are three options A, B and C, and A is indifferent to B and B is indifferent to C, then A is indifferent to C. If the rule applies for three options then it also applies for any greater number.

Dominance

If there are two options A and B which are assessed on a number of criteria, and A is preferred to B as assessed on one criterion while being preferred to or indifferent to B on each and all of the other individual criteria, then A is said to dominate B. It follows from this that A is preferred to B. This result also applies for any number of options.

Because these axioms are the bench-mark by which decisions are judged by the decision analysis community, the heuristic elimination strategies outlined in section 4.2.1 can be shown to be deficient because they can lead to intransitive choices (as defined by maximising methods) if unwisely applied (Montgomery, 1983; Ranyard, 1977).

4.4.2 The Measurement Problem

Having laid down axioms of consistent choice, there is then the problem of how to express preferences in such a way that the axioms may be adhered to. Typically, one of the criteria on which the decision will be based is cost, which is easily measured by money (although it might be argued that this is not so straightforward in some cases). However, performance criteria are not generally simple to measure in a way that enables consistent comparison with each other and with cost.

These criteria may be *incommensurable* or *intangible*. Incommensurable criteria are those which are measurable in their own terms (such as temperature, area, time) but which cannot be readily compared. Intangible criteria are those which have no natural unit of measure (such as aesthetic appeal).

The heuristic elimination strategies described in section 4.2.1 partly avoid this difficulty by comparing each criterion only in its own terms, however as it has been stated, such strategies are non-compensatory and can lead to intransitive decisions. Dealing with these measurement problems is therefore a fundamental requirement when wishing to evaluate options such that a maximising, compensatory decision strategy can be used. In essence, it is necessary to translate the performance on the various criteria into a common measuring denominator.

4.4.3 Cost-Benefit Analysis

One way to achieve a common measuring denominator is to attempt to translate the performance of the options on each of the criteria into monetary terms. This is what is done in Cost-Benefit Analysis (CBA) (Pearce and Nash, 1981; Sassone and Schaffer, 1978; Sugden and Williams, 1978; Layard, 1972). As originally developed, CBA is a technique which is intended to provide an objective measure of the economic effects of large social projects, however there have been suggestions in the past that the technique might be applied to more specific building design decisions (eg Warszawski, 1984; Fleming, 1973; Markus, 1973). Project costs are obviously already expressed in monetary terms, while the value of the benefits is measured by the market price of those benefits (ie what society is prepared to pay for them). If the costs and benefits of each option are accrued over an extended period of time, then a discount rate is usually applied to calculate the net present value of each option.

The technique is not without its problems, however. It can be extremely difficult to assess the market price of benefits, and even when assessed, their meaning can be questionable. Layard (1972) points out that in a mixed economy prices often fail to reflect properly the value which society places upon items, either because it is not truly a free market, or because the item is not a market good. When a direct market price for an item cannot be established, attempts may be made to find a 'shadow price'. This is the price for the item inferred by the societies willingness to pay for another item which the society shows itself to value equally (and for which a direct market price can be calculated). Often, however, there is no purely objective means of assessing shadow prices, and a degree of judgement must be used to adjust the best available market price indicators. Sassone and Schaffer (1978) warn that it is difficult to avoid arbitrariness when making such adjustments, which can compromise the objectivity of the CBA. Even with shadow pricing, however, there are many criteria involved in a decision which cannot be taken into consideration using the CBA technique, and it has been suggested that as a result these criteria tend to be ignored or at least under-emphasised (Watson, 1981).

The claim that CBA is objective and its reliance on market valuations has been criticised by several authors (eg Watson, 1981; Stern, 1976; Self, 1975). In a now famous paper, Self (1970) applies Bentham's epigram 'nonsense on stilts' to the cost-benefit analysis carried out by the Roskill Commission on the siting of the third London airport, and describes CBA as an attempt to "convert genuine political and social issues into bogus technical ones". In a similar attack on claims to objectivity, Stern (1976) suggests that "in reality, no-one can say what human life, health, quiet or even time are worth, and the assigned monetary values merely reflect the planner's prejudices".

The failure of market prices to reflect many aspects of a decision situation has lead to the incorporation of a form of weighted scoring analysis to supplement the purely economic assessment in some CBAs (Sugden and Williams, 1978). An early example of this is appendix 2 of the Buchanan Report 'Traffic in Towns' (Buchanan, 1963). Once again, however, the analysts purport to seek an objective means of arriving at the weights and scores used.

It follows from the foregoing, therefore, that although CBA is one way of dealing with the problem of common measuring denominator, there are technical and philosophical difficulties with the use of the technique which mean that it is not an obvious solution to the problem.

4.4.4 The Common Weighted Scoring Method

In another approach to the problem of converting performance on various incommensurable or intangible criteria into a common measuring denominator, the performance is converted into a numerical score on a form of points system. This is the basis for weighted scoring methods. The common approach to this as described in the design methods literature (Jones, 1970; Sanoff, 1977; Phal and Beitz, 1984; ASCE, 1988; Cross, 1989) may be described as follows:

1. First, all the decision criteria are listed.

2. These decision criteria are given a numerical weight according to their relative importance. (eg the most important criterion might be given a weight of 10, the second most important a weight of 7 because it is judged only 70% as important as the most important criterion, the next most important a weight of 5 and so on.)

3. The options are given a score on each criterion which reflects their performance. (eg an option which performs very well on a criterion might be given a top score of 10, while another option with a poorer performance might only be given a score of 2.)

4. The scores for each option are multiplied by the weights and these weighted scores are then added up to produce a total weighted score for each option. The option with the highest total score is then the one with the 'best' overall performance.

In mathematical terms, the combinatorial rules are a weighted sum model:

$$V_i = \sum_{j=1}^{n} w_j v_{ij} \tag{4.1}$$

Where V_i is the total weighed score for option i

w_j is the weight given to criterion j

v_{ij} is the score of option i on criterion j

This form of decision model can be criticised from both a theoretical and a normative perspective. In theoretical terms, Jones (1970, p380) warns that the arithmetic performed in this method "in ignorance of scaling operations can be as misleading as measurements taken with an elastic tape measure or as naive as the attempt to calculate the total size of an object by adding its weight to its volume". Even although a set of preferences expressed in terms of the simple weighted scoring method would conform to the axioms of consistent choice in section 4.4.1 above, it will be argued in chapter 5 that the method as it is presented here is deficient because of the lack of rigour in defining the set of criteria to be used in the evaluation and because of the combinatorial rules employed. It will also be shown, however, that a theoretically correct interpretation of the method (known as Multi-Attribute Value Theory or MAVT) is possible.

Those who seek objectivity (such as the proponents of CBA) have long further protested that, irrespective of theoretical arguments concerning criterion definitions and combinatorial rules, the criteria, weights and scores used in such an assessment are highly subjective (eg Beesley and Kain, 1964). This criticism of the method is unfounded, because as a 'soft' approach to decision making such methods make no claim to objectivity. What is being modelled is the *subjective value* which the decision actor places on each aspect of the decision situation. Accepting this, there remains the normative question of how to know that the criteria and the numbers supplied by the decision actor do represent a model which is an appropriate basis on which to evaluate the options. It will be argued in chapters 5 and 7 that generally these methods are also deficient in this respect, however it is possible to begin to address the failings by adopting the best practices advocated by the decision analysis community and by incorporating elicitation techniques developed in the field of psychology.

Chapter 5

A Critique of the Common Weighted Scoring Technique

Before proceeding to evaluate the appropriateness of weighted scoring methods to the task of selecting a generic HVAC system from a range of solution types, it is first possible to make a context-free assessment of certain aspects of the validity of the method.

This chapter presents a detailed critique of the weighted scoring technique as it is commonly described in the design methods literature and it is shown that the method is invalid by the standards of mainstream decision analysis. In developing this argument, the foundations of the theoretically correct weighted scoring method of Multi-Attribute Value Theory (MAVT) are laid down, and a succinct account of the theoretical basis for MAVT is given. The chapter concludes with the observation that the validity of MAVT in comparison with the common weighted scoring technique is in some ways a matter of degree rather than principle. Furthermore, it is highlighted that MAVT makes certain assumptions about the structure of the decision task which means that it is not a valid technique to apply in every situation that can be characterised as an appraisal of different solutions. Consequently, the notion of MAVT (or any maximising decision making strategy) as 'superior' to heuristic elimination strategies must also be qualified, and the context in which the technique is used must be examined carefully to ensure that the task is truly amenable to the application of MAVT.

5.1 THE DEFICIENCIES OF THE COMMON WEIGHTED SCORING TECHNIQUE

In this section, the deficiencies of the weighted scoring technique described in section 4.4.4 are discussed in detail. In the interests of clarity, the aspects of the method algorithm are considered in the order in which they are described in section 4.4.4, beginning with the definition of the criteria and concluding with the implications of the weighted sum aggregation rule. While reading the criticisms, however, it will be appreciated that these are not actually separate issues. The aggregation rule has implications for the definition of the criteria, the assessment of scores has implications for the assessment of weights and so on.

5.1.1 The Choice of Criteria

The structuring of the decision problem by selecting the criteria by which the options are to be evaluated is not given much attention in the design methods literature, although it is generally accepted that this is the most important stage in the entire process (Keeney and Raiffa, 1976; von Winterfeldt and Edwards, 1986). To demonstrate this, Keeney (1988) simply notes that the *de facto* weight of an omitted criterion must be zero.

The structuring process itself is seen to be more of an art than a science (Brownlow, 1989; Keeney, 1988) and undoubtedly relies on the creativity of the decision actor (and anyone assisting him), coupled with his particular view of the decision situation. In this respect then, many of the issues involved in the selection of criteria are extra-rational. However, notwithstanding this the end product of this structuring process has a tangible form which can, in principle, be well defined, and mainstream decision analysis holds that the chosen set of criteria should meet certain requirements if it is to constitute a theoretical and normatively correct model of the decision situation. Simple weighted scoring methods may be regarded as deficient because they fail to test whether the criterion set meets these requirements.

Keeney and Raiffa (1976) were among the first writers in decision analysis to attempt a definition of the desirable properties of a set of criteria. In terms of the theoretical and normative properties, they suggest that the set of criteria should be *complete, operational* and *non-redundant.*

A set of criteria is said to be *complete* if it includes all of the areas of concern in the problem at hand. For example, suppose the actor choosing a car from a range of options omitted to include luggage space as one of his criteria. He could find that he selects a car which, although the best overall on the other criteria which he did consider, is a regrettable choice when he tries to fit his golf bag and motorised caddy into the boot. Although completeness is a frequently sited requirement of the criterion set in the decision analysis literature, it may be argued that in fact this is an unattainable goal. All models are abstractions of reality and therefore by definition must be incomplete in some sense. On a more pragmatic level however, the notion that no important criteria should be omitted is acknowledged.

The criteria in the set are *operational* if they can be measured in a way that is meaningful to the decision actor, and which also facilitates explanations to others. In the motorcar example, the criterion 'power' could be treated as an all encompassing intangible statement of dynamic performance and measured on a semantic scale of 0 = poor to 10 = very good. This may be sufficient in some circumstances, but the danger with such scales is that the actor may be inconsistent in his judgement of what constitutes poor or very good, and he may also be inconsistent in the numerical translations which he makes from the underlying semantic scale. If he were to imagine attempting to explain the reason for his scores to a third party, he may find that relying on such a scale is inadequate. The applications of the simple weighted scoring method invariably use a semantic scale for all of the criteria with these attendant problems. In mainstream decision analysis, however, an attempt is made to avoid these difficulties by defining criteria such that they have a clearly understood natural scale of measurement whenever possible (ie are incommensurables rather than intangibles). Hence, the all encompassing intangible criterion 'power' might be re-defined or *characterised* by two new incommensurable criteria 'top speed' (measured in mph) and 'acceleration' (measured in the time taken to go from 0-60 mph). The consequence of this approach is generally to increase the number of criteria in the model in a hierarchical structure (ie with 'top speed' and 'acceleration' as sibling offshoot criteria of the parent 'power'). Although the number of criteria has no bearing on the theoretical properties of the model, in practice it has been observed that models with more than twenty or so lowest level criteria tend to become unwieldy and as a result the usefulness of the model in overcoming the actor's information processing limitations is impaired (Edwards and Newman, 1982). Inevitably then, a balance must be struck between the number of criteria and their operationalisation.

Non-redundancy requires that the set does not contain criteria which do not assist in differentiating among the options, or worse, effectively double count aspects of the decision situation. An example of a non differentiating criterion in the choice of a car might be 'all round visibility'. If all of the cars under consideration offer substantially the same degree of visibility, then the inclusion of this criterion adds nothing to the appraisal, but it does not corrupt it either. If the more serious error of double counting occurs, the effect is to increase the weight of the double counted aspect and distort the model. In an above example the criterion 'power' was re-defined as 'top speed' and 'acceleration', however it is also possible to include a third new aspect 'horsepower' (measured in bhp). If the actor is a frequent caravaner, and the ability of the car to pull loads is therefore important, then this third criterion would be a relevant and distinct aspect of 'power' which should rightly be taken into consideration. Unless this is so, however, 'horsepower' in itself might be irrelevant to the actor and its inclusion in the model would only result in the double counting of 'acceleration' related aspects.

In addition to these three requirements it should be added that the criteria must have a *range of convenience* which cover all of the options under consideration. This is a requirement which was formally highlighted by Humphreys *et al* (1980) who applied the psychology technique of Personal Construct Theory to structuring decision situations for Multi-Attribute Value Theory (a technique which will be considered in detail later). If a criterion is to be used in a weighted scoring type analysis, then it must have relevance to all of the options. For example if one of the options under consideration in the car selection problem is a convertible, then the criterion 'loading capacity of roof rack' would not have a *range of convenience* which covered all of the options because such a criterion is not relevant to a convertible.

5.1.2 The Meaning of Weight

This is the most contentious quantitative issue in the theoretical interpretation of weighted scoring methods (Keeney, 1988; Watson and Buede, 1987; von Winterfeldt and Edwards, 1986). In the common weighted scoring method, the decision actor ascribes weights to each criterion on the basis of their 'importance' such that high importance automatically implies a high weight. In a theoretical sense, however, mainstream decision analysis considers such an interpretation to be meaningless.

Suppose, for example, that an actor is asked to weight two criteria in the selection of a motorcar. He might say that cost has a weight of 10 and top speed has a weight of 2, ie cost is 'five times as important' as top speed. Would this mean that in order to save 10% on cost he is prepared to accept a car with a top speed reduced from 100 mph to 50 mph? Almost certainly not. The saving in cost does not justify the reduction in top speed. The point is that while the sentiment of a statement like 'cost is five times as important as top speed' makes sense, it is useless in making a decision between competing options. What is useful, however, is the notion of a trade-off between cost and top speed within certain reasonable limits. So for example, the same actor might say that while no reduction in cost would compensate for a reduction in top speed from 100 mph to 50 mph, he would be prepared to put up with a reduction of 5 mph to save 10% in cost. Mainstream decision analysis opinion holds this notion of a trade-off between criteria to be the correct interpretation of weight, and hence weights cannot be elicited from the decision actor without consideration of the range over which the options vary on the numerous criteria on which they are being assessed.

That this is the theoretically correct interpretation of weight can also be understood from the combinatorial rules of the weighted sum model.

From equation 4.1, the difference in total weighted score between two options V_1 and V_2 is:

$$V_1 - V_2 = \sum_{j=1}^{n} w_j v_{1j} - \sum_{j=1}^{n} w_j v_{2j}$$

$$= \sum_{j=1}^{n} w_j (v_1 - v_2)_j \qquad (5.1)$$

i.e. the weight w_j does not apply to the criterion j itself, but is actually a scaling factor for the intervals $(v_1 - v_2)$ on each criterion value scale j.

For example, take a case with two hypothetical options, $i \in \{1,2\}$ assessed on two criteria $j \in \{1,2\}$ which are indifferent to each other. It follows from equation 5.1 that:

$$V_1 - V_2 = \sum_{j=1}^{2} w_j (v_1 - v_2)_j$$

$$= w_1 (v_1 - v_2)_1 + w_2 (v_1 - v_2)_2 = 0$$

$$\Rightarrow \quad \frac{w_1}{w_2} (v_1 - v_2)_1 = - (v_1 - v_2)_2$$

i.e a difference of one unit on the criterion 1 value scale is equivalent to a difference of w_1/w_2 units on the criterion 2 value scale.

Notwithstanding the theoretical issues, several experiments have been conducted to determine whether the interpretation of weight makes a significant difference to the outcome of a decision, all other things being equal (Borcherding *et al*, 1991; Stillwell *et al*, 1987; Jaccard *et al*, 1986; Schoemaker and Waid, 1982). These experiments indicate that there is considerable variation in the values of weights elicited from the same subjects depending upon the interpretation of weight and the method of elicitation. Furthermore, although it is not possible to determine empirically which interpretation is 'correct' in that it leads to a better decision, Borcherding *et al* did attempt to approximate such an assessment by comparing how the decisions of a lay experimental group converged with those of an expert panel. They concluded that the theoretically correct trade-off interpretation is the most valid by this criterion.

5.1.3 The Meaning of Score

The score assigned to an option on each criterion can be interpreted as either a *ratio* or an *interval* measure. In the typical descriptions of the common weighted scoring method, however, the meaning of the score is rarely made explicit and the implications of these alternative interpretations are seldom explained.

A *ratio scale* is a scale with a natural zero. There are many examples of ratio scales in the physical world, such as length, time, mass, and so on. Such phenomena can be measured on a scale where the zero of the scale really does mean 'nothing' - no length, no time, no mass. Hence, irrespective of the measuring scale chosen, the zero is always the same eg zero centimetres = zero inches = zero fathoms. Furthermore, on such scales it is meaningful to speak of one measurement being twice or *n* times larger than another (eg one object can be twice or three times longer than another).

The alternative interpretation is that the score for each option is measured on an *interval scale*, ie a scale which is liner, but with a zero which is arbitrary. A common example of such a scale in the physical world is temperature, as measured on the Celsius or Fahrenheit scales, where zero does not mean 'no temperature'. Both scales measure the same phenomenon, but zero degrees Celsius is thirty-two degrees Fahrenheit. On interval scales it is meaningful to say that one thing is greater than another to a given extent (eg 40 °C is 10 °C higher than 30 °C, which is also 10 °C higher than 20 °C), but it is meaningless to speak of one measurement being twice or *n* times greater than another (eg 40 °C is *not* twice as hot as 20 °C).

The importance in the interpretation of the scale (from a theoretical point of view) lies in recognising the limitations on the form of arithmetic operations which may legally be performed using the scale. Ratio scales are the highest form of measurement scale and there is no limitation on the arithmetic operations, however the operations which may be performed using interval scales are more restricted. In general, operations may only meaningfully be performed on the *differences* between measurements on interval scales.

It is generally accepted in the decision analysis community that value scales are interval and *not* ratio scales (von Winterfeldt and Edwards, 1986; Dyer and Sarin, 1979). To return again to the example of the alternative motorcars, a decision actor might have said that on the criterion 'comfort' Car A scores 8, Car B scores 6 and Car C scores 2. This can be interpreted to mean that the perceived difference in comfort between Car A and Car B is only half that between Car B and Car C, but it cannot be said that Car A is 'four times' more comfortable than Car C, because the concept of zero comfort is not meaningful, even to those who have ridden in a Citroen 2CV. The zero point on such scales is arbitrary and may be defined as, for example, the poorest standard of comfort of any of the options, or as the poorest acceptable standard of comfort. These considerations will be discussed further in chapter 7.

5.1.4 The Meaning of the Total Weighted Score

If the individual criterion scores are measured on an interval scale, then it follows that the total weighted scores are also measured on an interval scale (weights are ratio concepts as they define trade-offs). Hence, it is the differences between the total weighted scores, rather than their absolute values, which are meaningful in terms of arithmetic operations.

Often, applications of the weighted scoring technique take the form of a performance versus cost analysis where performance is represented by a total weighted score and cost is naturally measured in money. The final stage of the analysis is then to find the option which gives the highest performance in relation to cost. Some published examples of such analyses (eg Gilleard, 1988; Buchanan, 1963) proceed to do this by finding the option with the highest ratio of performance to cost. This arithmetic operation is only valid given the assumption that the total weighted scores are measured on a ratio scale, however, and so in a theoretical sense such a calculation is meaningless.

5.1.5 The Assumption of Additive Difference Independence

An important assumption, and one which is never explained in the context of simple weighted scoring methods, is *additive difference independence*. For the weighted sum model (equation 4.1) to hold, it is assumed that the value of an option as assessed on any one criterion is not affected by the values of the other criteria. For example, suppose that motorcars are being assessed on the criteria cost, top speed and space, the latter two criteria being scored out of 10 and appropriately weighted. One car costs £10,000 and scores 8 for top speed and 5 for space. However, the supplier then increases his price to £11,000. If the additive difference assumption is to hold, this should not affect the scores of 8 and 5 awarded for the other two criteria nor their respective weights. In this example, it is reasonably safe to presume that the assumption holds, however there are instances where the assumption may not hold. Sometimes a poor performance on one criterion can have the effect of diminishing the perceived values on other criteria also. It is therefore necessary to check for such interactions to ensure that they do not exist. If they do, then the weighted sum model is not an appropriate aggregation technique. Formal proof of this result was originally developed by Fishburn (1965).

In terms of the mathematical model, the underlying assumption of additive difference independence may be illustrated as follows:

For example suppose there are four hypothetical options $i \in \{1, 2, 3, 4\}$ assessed on two criteria $j \in \{1,2\}$. Let the individual scores of options 1 and 2 as assessed on criterion 1 (v_{11} and v_{21}) be at a maximum (e.g 10) and let the individual scores of options 3 and 4 as assessed on criterion 1 (v_{31} and v_{41}) be at a minimum (e.g. 0). Further, let the individual score of option 1 as assessed on criterion 2 be greater than that of option 2 by the same amount as option 3 is greater than option 4 (i.e. $(v_1 - v_2)_2 = (v_3 - v_4)_2$).

It follows from equation 5.1 that:

$$V_1 - V_2 = w_1 (v_1 - v_2)_1 + w_2 (v_1 - v_2)_2$$

$$V_3 - V_4 = w_1 (v_3 - v_4)_1 + w_2 (v_3 - v_4)_2$$

but $(v_1 - v_2)_1 = 0$ and $(v_3 - v_4)_1 = 0$

and $(v_1 - v_2)_2 = (v_3 - v_4)_2$

therefore $V_1 - V_2 = V_3 - V_4$

i.e the fact that v_{11} and v_{21} are at maximum while v_{31} and v_{41} are at minimum does not affect the overall value judgement of the differences between V_1 and V_2 and between V_3 and V_4

5.2 MULTI-ATTRIBUTE VALUE THEORY (MAVT)

The seminal work on this subject is generally recognised as Keeney and Raiffa (1976), which draws together much of the work in the field of mainstream decision analysis. Another major contribution to the literature is von Winterfeldt and Edwards (1986). A full description of the development of MAVT such as is contained in these texts, most of it necessarily mathematical, is beyond the scope of this thesis. What follows in this section is a succinct account of the basis of MAVT, which builds upon the above critique of the simple weighted scoring method, and concludes with an observation on the claimed superiority for maximising decision strategies such as MAVT.

5.2.1 The Existence of Value Functions

Up till now, the notion that preferences can be formally modelled (as opposed to *expressed* which is another issue) in numerical terms has been taken for granted. In fact it can be argued that if an actor considers a (large) range of options, assesses them by a number of criteria (or attributes) and is subsequently able to produce a complete and consistent preference order over the options, then it follows that some form of value function *V* exists (Watson and Buede, 1987). The thrust of this argument is that given this complete preference order, conceivably a value function could be constructed by attaching a set of scores for each option on each criterion until one is found which reproduces the complete ordering of options.

Of course, in practice, the reason for using an appraisal method such as weighted scoring or MAVT is that actors find it difficult to produce such complete and consistent preference orders directly in complex decision situations, however the argument for the conceptual existence of some form of value function *V* remains valid.

5.2.2 The Form of Value Functions

The form of the value function *V* could be highly complex, depending upon the degree of interaction between the various criteria. In theory, the form could be so complex that the use of the function for decision analysis purposes would be impossible. Fortunately however, it has been found that in most cases preference structures exhibit additive difference independence, or at least do not violate it to any serious extent, hence the weighted sum form of value function in equation 4.1 can usually be applied without serious error (Fischer, 1979; Pitz *et al*, 1980; Hwang and Yoon, 1981; von Winterfeldt and Edwards, 1986). When unacceptable violations of additive difference independence do appear, it is usually possible to correct them by adjusting criterion ranges or re-defining the criteria (Watson and Buede, 1987; von Winterfeldt and Edwards, 1986). Specific tests for additive difference independence, and measures to correct violations are discussed in chapter 7.

Two other forms of value function also appear in the literature - the *multiplicative* form and the *multilinear* form - which permit slight relaxations of additive difference independence and allow for certain interactions between criteria (Keeney and Raiffa, 1976; von Winterfeldt and Edwards, 1986). In practice, however, the multiplicative form only permits such minor relaxations of additive difference independence that, as von Winterfeldt and Edwards put it (p 292) "it is difficult to think of convincing examples in which MDI [the multiplicative form] holds and ADI [the additive form] does not". The multilinear form is somewhat more general in the permitted criterion interactions, however this generality is accompanied by complexity which makes the form too unwieldy in practical applications. "In fact, when the number of attributes exceeds four, the number of necessary interaction parameters becomes mind boggling, and consequently the necessary assessment task becomes unfeasible" (von Winterfeldt and Edwards, 1986 p 293). Because of their limited practical application, these forms of the value function will not be discussed further.

Although no formal reasons for the general validity of the weighted sum model are given in the literature, it might tentatively be suggested that this fact is not too surprising given the context in which MAVT is normally used. Typically, it is applied to decisions involving the selection of an option from a short list of candidates. Any option with a performance on a criterion which is so poor as to effectively devalue its performance on other criteria will probably have been weeded out prior to the use of MAVT, leaving only those options which are *prima facie* contenders. It might be speculated that this action, coupled with the range of convenience requirement of the criteria, tends to reduce the criterion definitions and ranges to such an extent that additive difference independence either holds or holds approximately.

To extend the example of the motorcar given earlier (section 5.1.5), suppose that instead of increasing the price from £10,000 to £11,000 the supplier increased it to £20,000. This is an extremely large increase and it is unlikely that any improvement in top speed or space could adequately compensate for it, given a reasonable range of other options. However in all probability had £20,000 been the price initially, this car would not have appeared as an option on the short list in the first place.

5.2.3 The Weighted Sum MAVT Model

The foundations for the weighted sum MAVT model have already been laid by highlighting the deficiencies of the simple weighted scoring method. Essentially, it is a weighted scoring method which recognises the theoretical and normative interpretation of criteria, weights and scores.

The composition of the model is described in outline below:

1. The attributes/criteria j by which each option is to be judged are listed and tested for completeness, operationalisation, redundancy and range of convenience.

2. The performance of each option i on each attribute j is given an attribute value v_{ij} scored on an interval scale (eg 0 - 10). For incommensurable attributes with a natural underlying measure which will be denoted here by x_j, an *attribute value function* $v_j(x_j)$ is normally specified showing graphically how the value v of the underlying measure x varies throughout its domain.

3. The attributes are checked for additive difference independence, and re-definition of the attributes or re-examination of the range of options is carried out if a serious violation occurs.

4. The attributes are given weights w_j which reflect the relative trade-offs between a unit on the attribute value scales.

5./

5. The scores for each option are multiplied by the weights and these weighted scores are then added up to produce a total weighted score for each option. The option with the highest total score is then the one with the 'best' overall performance.

Specific methods by which criteria, weights and scores can be elicited, and by which additive difference independence can be verified are described in chapter 7.

5.3 CONCLUSIONS OF THE CRITIQUE

The foregoing critique has highlighted the deficiencies in the common weighted scoring technique when judged by the standards of mainstream decision analysis, and the previous chapter discussed the deficiencies of heuristic elimination strategies when compared with maximising strategies like MAVT. It is clear from this chapter, however, that the validity of MAVT in comparison with the common weighted scoring technique is in some ways a matter of degree rather than principle. Furthermore, MAVT makes certain assumptions about the structure of the decision task which means that it is not a valid technique to apply in every situation that can be characterised as an appraisal of different solutions. Consequently, the notion of MAVT as 'superior' to heuristic elimination strategies must also be qualified.

5.3.1 The Superiority of MAVT over the Common Weighted Scoring Technique

The definition of score, the definition of weight and the failure to verify the range of convenience requirement for the criteria are all clear points of theoretical principle where the common weighted scoring technique is demonstrably invalid. MAVT can be shown to be valid in these areas and therefore superior to common weighted scoring. There are, however, other areas in which validity over the common weighted scoring technique is merely a matter of degree, the achievement of which must be judged pragmatically. The extent to which a criterion set is complete, operational and non-redundant, and the verification of additive difference independence must be couched in relative terms. Weights and scores, while having clear theoretical definitions, must be elicited from decision actors with human frailties, in which case it would be unrealistic to expect complete accuracy in the numbers assigned. The validity of a decision model in these terms must therefore be a compromise between rational requirements on the one hand and practical constraints on the other.

Notwithstanding whether it be a matter of degree or theoretical principle however, the the simple weighted scoring technique can be shown to be invalid, and this must give cause for concern over the widespread advocacy of the method.

5.3.2 The Superiority of MAVT over Heuristic Elimination Strategies

The example of additive difference independence in section 5.2.2 above highlights how the application of maximising decision strategies such as MAVT depends upon the use of heuristic elimination strategies such as those discussed in section 4.2.1 in the first instance, and consequently the notion of them as 'superior' strategies must be qualified. The impression created by many writers (eg Sage, 1981; Janis and Mann, 1977; Etzioni, 1967) is that heuristic elimination strategies are short-cuts to decision making which are employed either consciously when it is not worth expending the effort on maximising, or unconsciously through a deficiency in the information processing of the actor. It is clear from the above, however, that these strategies are capable of dealing with decision situations which MAVT cannot cope with practically or, in some cases, even theoretically.

Firstly, the application of MAVT assumes that the relevant criteria can be defined in a set which has an appropriate range of convenience. Only certain forms of heuristic elimination share this requirement. Secondly, it is assumed that the range of options has been whittled down to a point where the attribute ranges are small enough for additive difference independence to hold or hold approximately. Thirdly, the requirement in MAVT to define upper and lower limits on the attribute value functions is in effect an application of the conjunctive decision strategy. Finally it can be demonstrated that, even in theory, no value function, *additive or otherwise*, can be constructed to model heuristic elimination strategies and thereby be used to perform this winnowing down process. This is because the threshold type behaviour rules out the existence of a continuous variable function (Brownlow, 1989) as can be proved thus:

Take the simplest case of a decision situation with two attributes $j \in \{1,2\}$. If a value function $V(v_1 , v_2)$ exists which can model lexicographic or elimination by aspects type heuristic elimination strategies on continuous attribute value variables v_j, then it must have the property that:

$$V(v_{11},v_{12}) > V(v_{21},v_{22}) \text{ only if } v_{11} > v_{21}, \text{ or } v_{11} = v_{21} \text{ and } v_{12} > v_{22} \quad (5.2)$$

where $i = 0, 1, 2 \ldots$ etc denotes the particular level v_{ij} of the attribute value variable v_j in a given range (e.g 0 - 10).

This implies that V must be a monotonic* function increasing in v_1 and v_2, which in turn implies that $V(v_1, v_{02})$ and $V(v_{01}, v_2)$ are also monotonic increasing in v_1 and v_2 for some fixed v_{02} and v_{01} respectively. Further more by the properties of increasing monotonic functions (see Kolmogorov and Formin 1970) it can be stated that there exists some level v_{01} for which $V(v_{01}, v_{02})$ is continuous in v_1 about v_{01}..

Now consider some level $v_{12} > v_{02}$

$$\rightarrow \; V(V_{01} , v_{12}) = V(v_{01} , v_{02}) + \alpha \text{ for some } \alpha > 0$$

However, continuity in v_1 about v_{01} requires that there exists some level v_{11} such that

$$V(v_{11} , v_{02}) - V(v_{01} , v_{02}) < \alpha/2 \text{ say}$$

This requirement for continuity conflicts with the requirements of equation 5.2 because although $v_{11} > v_{01}$ it is possible that $V(v_{11}, v_{02}) < V(v_{01}, v_{02})$

i.e No continuous function of the variables v_j exists which can model these heuristic elimination strategies.

* A *monotonic* function is continuous with only one value of the dependent variable for every value of the independent variable.

These limitations imply that although the preliminary study of chapter 3 found that the early stage HVAC design process typically begins with a two stage appraisal process, this cannot be taken to indicate that the weighted scoring technique of MAVT is valid when applied in this context. The following chapter reports on a second study which was carried out to determine whether the characteristics of the HVAC solution appraisal task are amenable to the application of MAVT in light of these limitations.

Chapter 6

A Study of the Early Stage HVAC Solution Appraisal Task

In the preliminary study reported in chapter 3, it was found that the process of HVAC design at the early stages typically begins with a two stage appraisal process in which the fundamental HVAC provision for each area of the building is assessed and then an evaluation is made of generic system types from within this provision category. This process is most intensive in the areas of the building considered to be critical, often culminating in a pro and con type analysis in which the advantages and dissadvantages of system types are weighed against each other. Furthermore, it was found that the most critical area for HVAC design in office buildings is the open plan/modular space, indicating that further research would most realistically be concentrated on this type of area. Despite the fact that this is a solution appraisal task, the previous chapter has shown that the limitations of weighted scoring techniques like MAVT mean that it cannot be assumed that MAVT is a valid technique to apply in this context. This chapter describes a further study which was carried out to determine whether the characteristics of the HVAC solution appraisal task are amenable to the application of MAVT in light of these limitations.

6.1 THE RESEARCH METHOD FOR THE SOLUTION APPRAISAL TASK STUDY

While the intention in this second study was still to build a partial model of the technical design task based on data, indicating the use once again of a grounded theory approach to analysis, the requirement to understand design decision making task in greater depth entailed a different approach to information gathering than that of the preliminary study. As was explained in chapter 3, the information gathering methodology of the preliminary study is not suited to uncovering the more detailed cognitive processes of designers. Although project files contain sufficient detail to provide information on the level of design development and the overt management process which guides this development, they do not contain information on the more fundamental underlying reasoning processes. Similarly, the use of standard interviews, while arguably adequate for uncovering the processes which were the focus of the preliminary study, are less suited to conducting research on more detailed aspects of the designer's reasoning processes. Research indicates that retrospective accounts of detailed cognitive processes which have become deeply internalised and virtually automatic through experience may in fact be post-hoc rationalisations (Watson and Evans, 1975) or personal theories (Nisbett and Wilson, 1977) rather than accurate accounts of the processes.

The problem of how to elicit models of reasoning processes which are as free of these corrupting influences as possible has been the subject of much research and debate recently, as understanding cognitive processes is a fundamental requirement for expert systems. A special area of interest within the expert systems field known as *knowledge elicitation* or *knowledge acquisition* emerged in the mid 1980s and there is now a large literature on the subject (see for example Hart, 1986; Kidd, 1987; Gruber, 1989; Firlej and Hellens, 1991). The construction of an expert system is not the object of this study, however it has been found that the techniques of knowledge elicitation practised in the expert system field are applicable to general tasks of knowledge transfer (for example, work on the use of knowledge elicitation techniques outwith the specific area of expert work on the use of knowledge elicitation techniques outwith the specific area of expert system development is on-going at the department of computer science, Brunel University under the direction of Nancy Johnson). Hence it was decided to learn from the experience of those in the knowledge elicitation research community and adopt appropriate methods from the range of techniques available.

6.1.1 Information Gathering using Knowledge Elicitation Techniques

Because retrospective accounts of detailed reasoning processes may be flawed, it is generally recommended that elicitation is done in the context of solving specific example problems (Hart, 1986; Waterman, 1986; Gruber 1989) which are often presented in the form of prototypical case vignettes. The use of vignettes can facilitate efficient elicitation interviews because they can be designed to acquire specific information and control for extraneous stimuli which are potential sources of confusion. Although they are necessarily highly succinct and contain information which is less rich than that which would be available in a real situation, it has been found that experts base decisions on only a few items of highly salient information. The knowledge elicitation research communtiy contend that provided the case vignettes contain this salient information, the expert's perception of the case as presented in vignette form will not differ significantly from his perception of a real situation. Notwithstanding these arguments, it is clear that the simplifications and artificiality which vignettes introduce may be a threat to the validity of the elicited information which must be assessed and accounted for.

The Prototypical Case Vignettes

For this study, four prototypical case vignettes were produced drawing upon the general knowledge of what constitutes the 'typical' form of design information for office buildings at the beginning of the feasibility stage gained from the preliminary study (see appendix 3). The vignettes are summaries of the information relevant to technical HVAC design, and they were devised to invite the potential use of the full range of HVAC services solutions for the open plan/modular office spaces from heating with natural ventilation through mechanical ventilation to comfort air conditioning (this study was not concerned with the HVAC provision for other spaces). Two of the engineers from the first study (who were not to participate in this second study) were asked to review the vignettes for realism and completeness, so far as this is possible in such circumstances, and to give an opinion on whether the full range of HVAC solutions could be applied.

In line with the task orientation of this research, the vignettes concentrated on the technical perspective of the decision situation, and ignored other factors such as the political background. So for example, the engineers would be free to make decisions based only upon the apparent technical requirements of the case and not have to consider complicating environmental factors such as 'bee in the bonnet' client fetishes or restrictive developer rules. Nor would they need to consider pressures which might arise from organisational interests like pushing up the design fees or experimenting with an HVAC solution the consultancy practice has not tried before. The selection of a particular perspective or *domain view* as it is sometimes called (Firlej and Hellens, 1991) is also a common technique in knowledge elicitation, where it is used to reduce the task of elicitation to manageable proportions.

In spite of the attendant artificiality, it was believed that the method of elicitation via case vignettes would not be so far removed from practice as to render the results unrepresentative of the real technical design decision making strategies adopted by the engineers. The preliminary study had shown that the early stage design of HVAC systems is a relatively self contained process, where the engineer relies mainly on the information before him and makes safe assumptions about the things he does not know (ie temporarily treats the ill-structured problem as well-structured so that a solution can be attempted). If an engineer indicated that consultation with others would be normal at a certain point, then this fact could be noted. The problem would lie in the absence of feedback from such consultations. An attempt had been made to anticipate most of the questions which an engineer might ask of a client or architect (hence the inclusion in the vignettes of information such as room for negotiation in slab to slab heights which the engineer would normally need to wrest from the architect) and so the artificial isolation from a client and a design team was not thought to be a serious problem, but a situation which "can be lived with" as one of the reviewing engineers put it. With regard to the absence of political pressures, the engineers who reviewed the vignettes indicated that it was an ideal sort of situation which would permit a truly conscientious approach to design in the client interest, and was not too unrealistic in terms of better quality owner occupied buildings (which is why the case vignettes were designated as such). Nevertheless, while these are *a priori* arguments for the adequacy of the information gathering method, checks on the methodological validity of the study were obviously required and these were incorporated into the interview and analysis procedures which are described below.

The Interview Procedure

Four of the eight project engineers from the preliminary study participated in this second study. At all stages of the study the interviews were conducted separately at their consulting offices.

Firstly, an informal *orientation discussion* (Hart, 1986; Mittal and Dym, 1985) was conducted in which the aims of the study were explained. The engineers were informed as to the assumptions which were to be made regarding the domain view, and the level of detail at which their reasoning processes were to be modelled. It was explained that the processes of interest are those which guide the selection of a generic system type at the earliest stages, and not the calculations and so on associated with 'drawing board' design.

The actual elicitation interviews were conducted in three sessions. In the first and second interviews, all four engineers were presented with the same four prototypical case vignettes and asked to talk through the reasoning by which they would decide upon the most suitable generic HVAC system type. A reasoning process elicited using one vignette was then tested for its transferability to the others by teaching back (see below) the process in the context of other vignettes. In this way informative similarities and differences in the cases were highlighted, and a greater degree of generality in the reasoning processes was achieved.

In spite of the greater degree of structure which this method introduces, the research (in common with most elicitation exercises) was still conducted in an exploratory spirit. While the vignettes were the *context* for the elicitation, it was not necessary to adhere to the letter of each one and hypothetical variations on them were also discussed. Hence, 'what if' questions were asked, such as 'what if the plan depth had been 18 metres instead of 12 metres?' or 'what if it had been specified that the temperature had not to exceed 26 °C?'. Additional questions of this kind which related to the sensitivity of certain parameters, but which had not been anticipated, were also raised in response to points made by the engineers. For example, two engineers suggested that the suitability of natural ventilation in vignette 2 would depend upon wind conditions on the site (which were not specified in the vignette). The other two engineers had not raised this issue, perhaps because of assumptions they had made or because nothing in the vignette had triggered them into considering this point. In a subsequent interview a non-leading queue in the context of a teachback response (see below) was then used to trigger this consideration on the part of the other engineers '...and there are *no* circumstances in which natural ventilation *can't* be used as long as the site isn't too poluted or noisey...'.

In the third and final elicitation interview, the process which had been induced from the previous interview data was summarised and the engineers were asked for comments. In particular, they were asked to think of any case examples from their experience which might introduce a difference of any kind to the process which had been based only on the prototypical vignettes.

This precaution was thought necessary because while the selection of a domain view and the use of vignettes enhances the reliability of the study (see section 3.1 for a definition of the four aspects of methodological validity; namely construct validity, internal validity, external validity and reliability) it places restrictions upon external validity. It is possible that the process elicited might be in some ways an artefact of the elicitation methodology and therefore not properly representative of the general reasoning process. In the event, it was found that the exceptional cases did not alter the model to any significant extent, but merely introduced two modification loops which are described in section 6.2.2. It should also be noted that the object of this study was not to build a generally applicable model in expert system fashion, but merely to explore the *form* of decision making task as perceived by the engineers.

Construct validity is a major concern in knowledge elicitation. Being sure that the interviewer has correctly understood the interviewee is extremely important when there is no other form of data available against which the perceptions may be tested or 'triangulated'. Most writers in the field of knowledge elicitation caution most strongly on this subject, and some enumerate the various traps into which interviewers might fall. The most insidious threat to construct validity is the fact that the interviewer and interviewee may have different interpretations of the same substantive information. That is to say that, in both a metaphorical and literal sense, they may not 'see the same thing'. Firlej and Hellens (1991) point out that this kind of misunderstanding is quite common in learning contexts, and part of the role of a teacher is to correct a student's misconceptions. However, they go on to say that "in the case of knowledge elicitation, there is no 'teacher' role as such, the expert and elicitor are engaged more in a mutual exploration because the expert most likely does not have a conscious model of the knowledge. The dangers of the elicitor misunderstanding in effect, are therefore that much greater" (p72). In spite of the importance of this issue, however, few writers actually suggest specific practical techniques by which the problem of misunderstanding can be minimised *during* the elicitation interviews. The consequence of this is that misunderstandings will only become apparent at a later stage when (in the context of expert systems) an early version of the system is tested out (Waterman, 1986).

One practical technique by which construct validity can be checked during elicitation is *teachback interviewing* which was first applied in the field of knowledge elicitation by Johnson and Johnson (1987). Teachback interviewing has its fundamental basis in conversation theory which is, in its entirety, a complex psychological model of the exchange of information between parties. On a practical level, however, the application of teachback interviewing in the context of knowledge elicitation is quite straightforward. Pask (1974) contains the full original description of pure conversation theory, and Obgorn and Johnson (1984) is a more succinct account of how the theory can be applied in knowledge elicitation.

The part of conversation theory which concerns construct validity in knowledge elicitation deals with the notions of concepts and understanding as made public by an interaction between participants. The analysis of an interaction can be considered at two levels: Level 0 (L_0) and Level 1 (L_1). Level 0 procedures bring about the preliminary specification of a concept and a description of reasoning algorithms. For example, an engineer might be asked 'how did you decide that natural ventilation could be used in this case?', and he may describe, in loose algorithm fashion, his reasoning procedure, which is also indicative of his concept of 'natural ventilation'. At Level 1, there are the methods for the reconstruction of L_0 procedures, ie how one knows how to perform L_0 procedures, and why one can be sure that they work. So, for example, an engineer might be asked 'on what basis were you able to decide that the building plan was not too deep for general natural ventilation?' (if this judgement is implied by his previous answer), and he may explain that experience had taught him that natural ventilation would be acceptable given a plan depth of no more that *x* metres, or that certain codes of practice showed that natural ventilation would work up to a plan depth of *y* metres.

Johnson and Johnson summarise the practical application of teachback interviewing as follows:

> "In a conversation E [the expert] will describe a procedure to A [the analyst]. A will teach it back to E in E's terms and to E's satisfaction. When they agree that A is doing the procedure E's way, then it can be said that A and E share the same concept. This 'teachback' procedure is a checking device where E is the final judge. Notice that A and E do not necessarily have the same private thought processes. They just agree that the same thing has been done. Then A asks E to give an L_1 explanation of how he reconstructed that concept and the teachback procedure continues until E is satisfied with A's version. Then we can say that A has understood E. Briefly then: at L_0 procedures define concepts; at L_1 reconstructions define 'memories', which, through teachback, lead to understanding."

The technique of teachback interviewing introduces the role of teacher and student to the elicitation process, but in an informal way and with reversible roles which preserves the spirit of mutual exploration and allows a formalised model of the expert's knowledge to be 'argued out'. In this study, the main concern is with the form of decision making strategy used by the engineers, and so for the most part it was sufficient to confine the elicitation to L_0, with only a general understanding of L_1 processes required to place the L_0 processes in context.

6.1.2 Analysis using Grounded Theory

During and following the elicitation process, normal practice in the field of expert systems development is to construct a formal model of the expert's knowledge in preparation for encoding into some form of knowledge representation format (Waterman, 1986; Liebowitz and De Salvo, 1989). Some techniques exist which can assist in this process by finding associations between items of information, however in the main the work of model construction is left *ad hoc* to the ingenuity and creativity of the elicitor (Trimble, 1988). In this sense, the inductive model building task is similar to that of chapter 3, and once again the discipline of grounded theory was applied as an aid to analysing the data gathered.

The aim here then, was to examine the data collected in the information gathering interviews and allow a grounded theory model of the solution appraisal task as manifested by the engineers' decision making strategies to emerge from patterns decerned in the data. The use of grounded theory enhanced the internal validity of the research findings because the patterns from which the decision strategy model was constructed emerged from the data, and because the patterns could be matched across each of the engineers. As an additional check on both internal and external validity, the decision strategy models were also verified against the data from the preliminary study, where the information was gathered using more naturalistic methods. Although the data regarding specific decision strategies was considerably less detailed, the more overt behaviour of the engineers recorded at that stage is presumably a higher level manifestation of these decision strategies, and so should be compatible with them.

6.2 PATTERNS WHICH EMERGED FROM THE DATA

The analysis of the data using grounded theory established patterns in the information elicited from the engineers. The orientation interviews revealed that the engineers share a perception of current issues in HVAC design which to some degree shapes the approach of the engineers to the HVAC selection task. It was also confirmed that the process of selecting a generic HVAC system type comprises two distinct stages. First, the fundamental HVAC requirements for the space are determined from among a range of broad generic provision categories, then an assessment is made of generic system types from within this provision category. The patterns found enabled the construction of models which represent the decision making strategies of the engineers at both stages.

6.2.1 Current Issues in HVAC Design

The orientation interviews revealed current issues in HVAC design in the UK which influence the engineers' decision strategy. As it was helpful to be aware of these issues when interpreting the data from the elicitation interviews, these issues are summarised here.

General Perceptions

* Air conditioned buildings are higher users of energy and the refrigerants associated with them are damaging to the environment. From a general environmental perspective, natural ventilation is preferred to mechanical ventilation which in turn is preferred to air conditioning, and the Building Research Establishment Environmental Assessment Method (BREEAM) reinforces this view.

* Almost all so-called sick buildings are air conditioned buildings. Furthermore, the general impression at large is that the causes of sick building syndrome are a mystery, implying that even good design practices cannot guard against it. This has made some clients wary of air conditioning. However, although it is not yet possible to explain all cases of sick building syndrome, the engineers believe that it is misleading to suggest that unidentifiable causes are predominantly to blame. Rather, they maintain that the majority of cases of sick building syndrome can be attributed to fairly obvious points of poor design and maintenance of the air conditioning system or of the building itself, and therefore can be controlled for in better design.

Technical Developments

* Recent research by the Department of Environment and others has shown that the practice of assessing the heat gains from office equipment by adding up the rating plate wattages has led to a three to five fold over estimation of power density.

* The thought that temperatures in office spaces should be maintained below about 23 °C is being questioned as people in naturally ventilated offices seem content to tolerate temperatures up to 28 °C on occasions.

* There is now available a range of computer packages which give designers the freedom to simulate the thermal performance of relatively novel building designs.

The engineers believe that these technical developments coupled with the general perceptions about air conditioning above have encouraged HVAC design engineers to question the need for air conditioning in smaller and medium sized office buildings where the heat gains are moderate and where higher temperatures for certain periods in the summer are acceptable. In particular, for deeper plan buildings, there is a trend towards experimentation with passive temperature suppression methods like exposed structural elements and external solar shades coupled with the use of high volume mechanical ventilation systems. It is possible that such measures could redude peak summertime temperatures to an acceptable (albeit high) level in some buildings, so removing the need for air conditioning. It is noted that the potential for such passive measures rests largely with the architectural design of the building. Although architectural modifications can sometimes be made to accommodate passive measures if the original building design is not suitable, HVAC considerations will not generally dictate architectural design.

There are, however, reservations about such developments. The engineers are concerned that the success of these passive design ideas depends in large part on the thermal simulation modelling used to test the design 'on paper', and they are aware that the computer packages available have not been fully validated because of the practical difficulties involved in validation exercises. It is believed that, at present, it is a brave designer who would opt for the passive solution if confronted with a borderline case.

6.2.2 Determining Fundamental HVAC Requirements

When presented with the sketched plan and data of a prototypical case vignette, the engineers first looked for major queues which would enable them to home in on the fundamental form of HVAC provision which was appropriate. It was found that the domain is relatively well structured, with the engineers adopting an essentially similar form of decision strategy outlined below. It should be noted that this does not mean they will reach the same decision in every case. Different interpretations of case study data and differences of opinion on what can be achieved in design terms were observed, and this could affect a particular decision outcome.

Is Comfort Air Conditioning Required?

If a particular temperature and/or humidity control band is specified then, irrespective of the type of building, the engineers judge that air conditioning is necessary. There is no other way of ensuring that temperatures will not rise above 23 or 24 deg C in hot weather. If there is no specific control band, however, the engineers will then assess whether the heat gains to the space are likely to be such that ventilation alone could suppress peak temperatures in the space to acceptable levels. This assessment usually involves a form of approximate thermal computer modelling in the first instance which will predict the peak summertime temperature in the space. If this assessment indicates that the peak temperature is too high, even given large ventilation rates, then air conditioning is once again indicated. If, on the other hand, it seems that the temperatures could be maintained at an acceptable level with ventilation only, the engineers then proceed to conduct a more detailed assessment of the required ventilation rates. In borderline cases, it may be that a building could be modified to make use of passive temperature suppressing measures such as exposed structural elements and external solar shades which would allow high volume ventilation to be used in stead of air conditioning. (As a rule of thumb it is believed that such measures can reduce the peak temperature in a space by around 3 °C).

Is High Volume Ventilation Required?

When high volume ventilation is being considered, the engineers conduct more detailed assessments of the building thermal performance using dynamic computer simulation packages. These simulations are considered particularly important where passive measures are also being used. If these assessments indicate that high volume ventilation will contain peak temperatures to within acceptable levels with a degree of safety (because the simulations are not accurate predictions), and if the building has sufficient space to accommodate the required duct work for such a system (which can be larger than for some air conditioning systems), then HVAC design will proceed on this basis. However, if it should transpire that the engineers are not satisfied that high volume ventilation can cope with the heat gains, or if they do not believe that the building could accommodate the ductwork, then air conditioning is indicated again.

Is Standard Ventilation Satisfactory?

If the heat gains to the space are such that overheating will not occur at standard ventilation rates, then the object becomes the provision of fresh air at the required rate. If conditions are suitable, then the obvious and preferred method is via opening windows (usually with ventilators for use in winter). However, various factors can preclude the use of natural ventilation via windows in certain circumstances, ranging from the obvious situation where spaces are isolated from such a window (which precludes its use in those particular spaces but not necessarily at the perimeter) to unfavourable site conditions (which precludes its use in any space). In these cases, a standard mechanical ventilation system is indicated, or possibly ventilation via an open atrium in certain buildings.

The requirement to provide heating does not affect the above considerations. It is assumed that heating will be needed at the perimeter of the building, but decisions on the form of system will only follow the outcome of the decisions on air conditioning and ventilation.

Figure 6.1 illustrates the above decision strategy graphically. As previously stated the particular domain seems to be well structured and the basic decision strategy depicted here is common to all. Where individual differences occur is in the threasholds at which the outcome of some condition tests (indicated by the diamonds) will be yes or no (for example, what is considered to be an 'unacceptable peak temperature' or a 'suitably ventilated open atrium' varies from one engineer to another). At this point the research was mainly concerned with discovering the *form* of strategy adopted by the engineers, and so a full examination of these differences was outside the scope of this study. Individual differences which did affect the strategy model, although in a minor sense, are the modification loops associated with high volume ventilation which were added as a result of the third interview in which the engineers were asked to think of any cases which did not conform to the pattern developed using the vignettes. One engineer said that he knew of a building which required high volume ventilation as originally designed (albeit a borderline example), but which was modified by the use of passive measures so that only standard ventilation was necessary. Another engineer commented that even although a building may not be able to accommodate a high volume ventilation system as originally designed, the possibility that its use would remove the need for air conditioning could encourage the client and the architect to reconsider the building design.

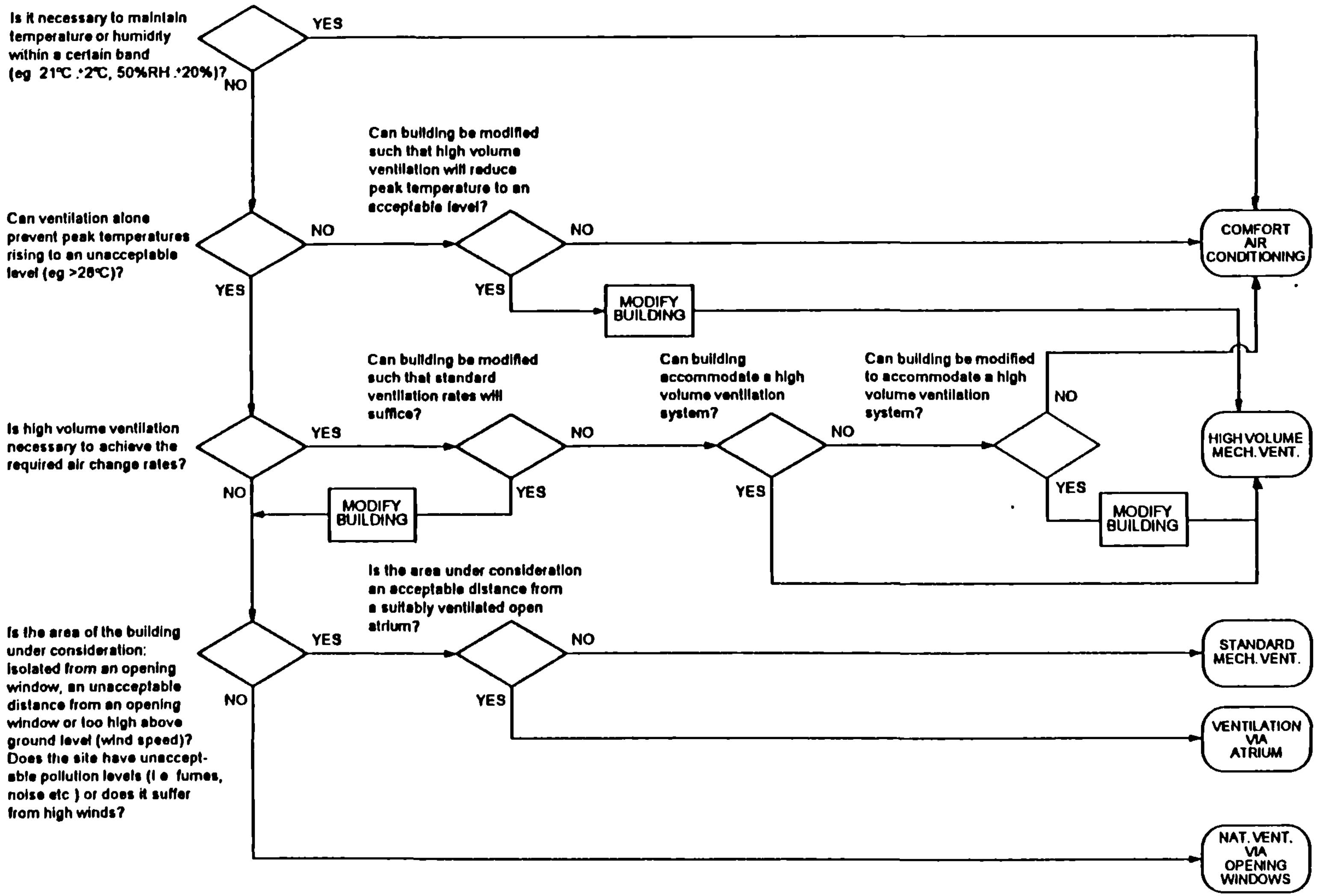


Figure 6.1 *A descriptive model of the decision strategy adopted by HVAC engineers when deciding upon the fundamental HVAC provision in an office space*

6.2.3 The Choice of Generic System Type

Once the engineers have decided upon the fundamental HVAC provision in a space, the problem then is to choose the type of system which would best meet the requirements. This problem applies to some extent no matter which fundamental solution has been decided upon, but is particularly difficult in terms of air conditioning because of the variety of generic types available, and because of the number of factors which have to be taken into consideration. One engineer commented "There are only really a couple of ways you can design a fresh air [mechanical] ventilation system and a heating system as far as generic types go...I don't think I'd spend long thinking about that...I'd be into what you [the interviewer] call drawing board design more or less straight away...The same with these high volume systems too, only they're more difficult...Air conditioning is different though. You've got a range of types and a few of them would probably do the job, although they've obviously got their advantages and disadvantages..." Another comment which was typical of observations on air conditioning system selection was made by another engineer: "If you look around you'll find different systems in more or less the same type of building, and they all work barring some disasters...When it comes down to it there might be a couple of things that rule out this system or that system, you know, not enough of a duct zone for VAV, too much refrigerant pipework for VRV...but apart from things like that three or four different systems could be made to work in most buildings."

It was decided at this point to concentrate on air conditioning system selection for the remainder of the study, as it appeared that this presented the greatest pre drawing board design challenge.

The actual basis for choosing a system type involves weighing up the advantages and disadvantages (or pros and cons) of each system on a number of criteria (or 'factors' as was the common term used) in the light of the particular building design and client requirements. The engineers were asked to think about the technical factors (criteria) they would consider when choosing a generic air conditioning system type. The object of this exercise was not to compile an exhaustive list, but merely to get an indication of what the engineers ment by 'factors'. The criteria are listed in table 6.1 with the number of engineers who mentioned each particular criterion. They clearly perceived some criteria to be 'more important' than others, and they demonstrated an awareness that the degree of importance may vary from case to case. All four engineers considered this assessment process to be a difficult problem because of the number of criteria involved and because of the absence of generally accepted rules. As one of them put it, "a lot of it is matter of opinion". This decision strategy is depicted in figure 6.2.

TECHNICAL CRITERIA	NUMBER MENTIONING EACH CRITERION
Capital cost	4
Running cost	4
Floor space	4
Services zone	4
Flexibility	4
Maintenance	4
Noise	4
Fresh air provision	3
Humidity control	3
Temperature control	2
Air distribution	2
Hygiene	2
Water damage risk	1
Length of refrigerant pipework	1

Table 6.1 The criteria which the engineers stated that they take into consideration when appraising different air conditioning systems

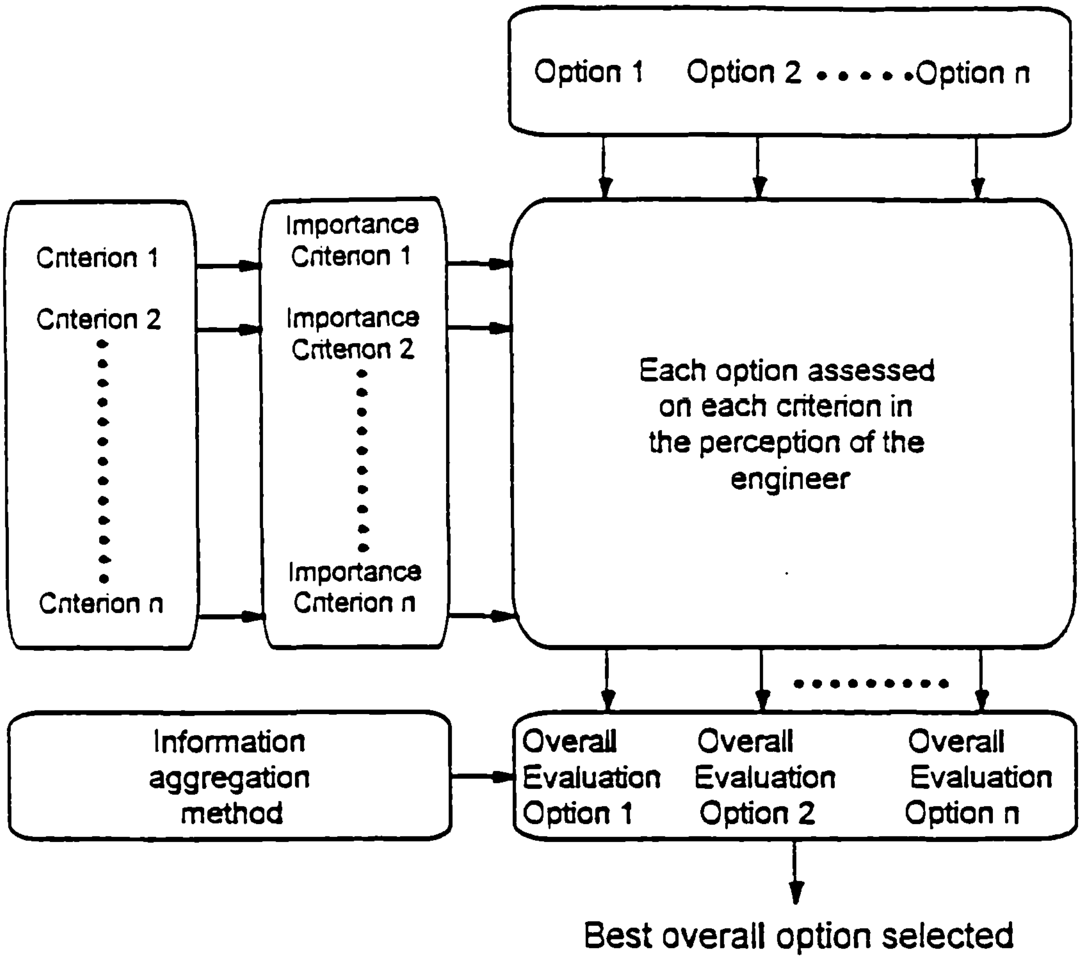


Figure 6.2 *A descriptive model of the decision strategy adopted by HVAC engineers when deciding upon the type of air conditioning system for an office space*

6.3 CONCLUSIONS OF THE DESIGN DECISION MAKING STUDY

It is clear from figure 6.1 that, in the initial stages, the decision strategy leading to the choice of fundamental HVAC provision is a process of heuristic elimination supplemented by additional procedures which allow modifications of the original building design. Subject to certain thresholds, (eg the existence of a specified temperature band; the need for high volume or standard ventilation) the engineer eliminates certain options until a final selection is made.

In this instance, it does not appear as if the heuristic process is a short cut to decision making. Save for the (limited) scope for modifications to the building design, each of the fundamental HVAC solution types are answers to quite different problems, except in the most trivial sense that they all provide a form of ventilation and temperature control. It would not be logical to attempt to build an MAVT model (or any maximising model) to select between air conditioning, high volume mechanical ventilation, standard mechanical ventilation and natural ventilation in a given building. Clearly not all of these options would work in the same building, and it would therefore not be possible to find a meaningful set of criteria with a range of convenience which covers all of the options.

The selection of a generic system type within the major provision categories, on the other hand, would appear to be characterised by a maximising type approach, albeit very informal. This attempt to maximise is most distinct within the air conditioning category (figure 6.2). There will usually be a number of air conditioning systems which are potential solutions to the same problem, and the selection is seen by the engineers as the identification of the 'best' system as appraised on a number of criteria.

Therefore, this study indicates that the task of determining the fundamental HVAC system provision is *not* amenable to a weighting and scoring technique like MAVT because the decision situation is not appropriately structured. However there is *prima facie* reason to believe that the more specific task of selecting a system type within the major provision categories may be an appropriate problem for MAVT, with further research being most beneficially aimed at air conditioning system selection.

Chapter 7

Research Method for the MAVT Validity Study

It has already been explained in chapter 4 that assessing the validity of a prescriptive decision aid is a difficult problem. Nevertheless, it is important that efforts are made to establish the validity of such aids. Given that there is now a clear decision situation with which an application of the weighted scoring technique of MAVT would seem, *prima facie*, compatible, this chapter describes the research method which was derived and subsequently adopted to make practical tests of normative and theoretical validity. Chapter 8 following details the application of the method in the third and final study of the research. Owing to the nature of the subject, it was known that it would not be possible to produce definitive answers. It was hoped, however, that some steps may be taken towards a useful assessment of validity in the context of air conditioning systems appraisal at least.

7.1 THE GENERAL APPROACH TO THE STUDY

In this section, the approach to studies in decision analysis carried out by the decision analysis research community is briefly reviewed and the approach required for this study is outlined. There is a paucity of work on the detailed validation of particular decision analysis methods, and so in this respect the research was venturing into uncharted waters. The intention was to test validity in a variety of different ways, based on the philosophy espoused by Watson and Buede (section 4.3.3). It was understood from the outset that the results of any such individual study would be inconclusive, however individual studies contribute to a 'web of evidence' which will ultimately indicate the general validity or otherwise of a method.

7.1.1 Methods in Decision Analysis Research

Investigations into decision aids can be pursued in a range of environments between the extremes which are analogous to the laboratory and the field (Belton, 1986). At the latter end of the spectrum, the research could comprise the application of a decision aid in 'live' situations as a form of action research. This was how Checkland and his team at the University of Lancaster developed the approach to problem solving which they termed 'soft systems methodology' (Checkland, 1981; Checkland and Scholes, 1990), and there are also numerous examples of research oriented applications of decision aids from within the mainstream decision analysis movement (Corner and Kirkwood, 1991). This research method can be a valuable source of information on the in-use performance of a decision aid, which is after all its ultimate purpose, however the many uncontrolled factors at play in a real situation can make the effects of the aid itself difficult to discriminate.

At the other extreme, laboratory experiments need involve neither a real problem nor a decision actor. It is possible, for example, to compare the output of different decision aids in theory only by using example problems and computer simulated inputs. This approach gives the researcher complete control over the situation, but is limited to highlighting differences in algorithms. Subject based experiments take a step away from this controlled environment and usually involve giving a number of actors a realistic but simplified decision problem and asking them to use a decision aid or several decision aids to solve the problem.

While retaining a degree of control over the situation, Hobbs (1986) points out that these so-called 'experiments' are nevertheless ill-controlled by the standards of experiments in the natural sciences. To note some of the confounding influences: the subjects themselves will vary in their perception of a decision problem, their disposition towards decision aids and their attitude towards the experimental situation in general; the presentation of the decision aid, while irrelevant in many senses, can have a distorting effect; and the decision analyst, if one is present, is a further confounding element. In spite of the attendant difficulties, however, subject based studies are indispensable because "(1) some knowledge is better than none, (2) experiments are a way of recording the history of methodological issues, (3) experimental results discourage useless metaphysical speculation and (4) experiments have an important effect on practice by showing what works." (Hobbs, 1986).

By far the most common form of subject based study carried out in decision analysis research is that which compares different decision aids applied by the same actors to the same problem (for important reviews and examples of such studies see Hobbs *et al*, 1992; Sainfort *et al*, 1990; Hobbs, 1986; Belton, 1986; von Winterfeldt and Edwards, 1986 and Hogarth, 1982). Studies of this kind are useful in demonstrating the practical and theoretical differences between decision aids, however they are limited in assessing validity. It is now well established that different methods can produce different results in that they can lead to different rank orderings of options. The difficulty is in knowing which one is 'right'. An assessment of validity requires a more in-depth investigation of a particular decision aid, and it is this problem that had to be addressed in the present study.

7.1.2 Approach Adopted For the Validity Study

Given that the focus of this research was the validity of MAVT as applied to a particular technical task, it was decided to adopt a subject based methodology similar to that often used in decision analysis 'experiments' for this study, and to adapt it to the specific aim of assessing the validity of MAVT applied to the selection of generic air conditioning system types for the open plan/modular spaces in office buildings. This would require presenting HVAC engineers with one or more example decision problems and using MAVT to assess several generic air conditioning system types in the context of the examples, making an assessment of the validity of the MAVT method during the exercise. It will be readily recognised that this presents several methodological challenges (see section 3.1 for a definition of the four aspects of methodological validity in research). Unfortunately, such is the nature of the subject under study that while steps may be taken to reduce threats to research validity, they cannot be eliminated.

First, external validity requires that the results of the study can be generalised to a degree. This calls for the use of as many subject engineers and as many example case vignettes as possible. However, addressing other methodological considerations discussed below places a practical limit on the number of subjects and case vignettes which could be handled. It was decided that six subject engineers each presented with three case vignettes would be the maximum number which could be managed. This was considered satisfactory in terms of analytical generalisation using the same replication logic as the previous two studies.

Second, case examples must be devised that are sufficiently realistic to satisfy external validity, while at the same time being simple enough for the purposes of the study. This problem had already been addressed in the decision making study of chapter 6, where it was found that case vignettes of the form used were reasonable representations of the real technical design task. The vignettes used for this study are discussed in section 7.2.

Third, there is the study situation and the confounding influences of the decision actor's disposition and the presentation of the MAVT technique. Few studies in decision analysis have been conducted to investigate these influences specifically, however the curious dispersal of results in comparative experiments such as those referred to in the previous section, and the tendency for repeated experiments to produce contradictory results (in terms of the perceived difficulty of use, transparency and so on of different aids as reported by the experimental subjects) is indirect evidence that factors such as actor disposition, aptitude with the method and the presentation of the method are important variables which are threats to construct validity and reliability (Hobbs 1992, Delquie, 1990).

Fourth, there is the application of the MAVT technique itself. As has already been pointed out, prescriptive methods are weak methods, because using them requires knowledgable interpretation of the various rules and terms, and, even if properly applied, success is not guaranteed. A consequence of this is that any study of the validity of a prescriptive method could be criticised on the grounds that the method was not properly applied - and such a criticism cannot be empirically refuted.

In answer to these third and fourth points, the most that can be done is to ensure that the best available practices have been used in the application of the MAVT technique, and that the subjects see beyond superficial aspects of technique presentation and understand the implications of the value model they are constructing. To this end, it was decided that the researcher would act as an analyst for the subjects to ensure that good practices were followed and that the subjects properly understood the method under investigation. The inclusion of an analyst introduces some potential facilitator effects which must be assessed and accounted for however. Issues of what constitutes good practice and possible facilitator effects are discussed in sections 7.3 and 7.4.

Finally, making an assessment of the validity of the MAVT technique requires developing more explicit statements of criteria whereby theoretical and normative validity of the technique may be recognised, thereby enhancing the construct validity of the research method, while at the same time guarding against threats to internal validity. The most pervasive and paradoxical issue in this regard is *forced consistency.* The issue is pervasive because the essence of rationality is consistency, and many of the assessments of technique validity must, in one way or another, rely upon the consistency of the MAVT model. It is paradoxical because it is at once both a strength and a weakness of a decision aid. Decision analysts will point out that actors faced with complex appraisals will naturally tend to make individual judgements which are inconsistent during the elicitation of weights and scores, and it is the job of the analyst (or the 'front end' of the decision aid) to highlight these inconsistencies and provide an opportunity for the actor to modify his judgements in line with the requirements for consistency in the model. It has been suggested, however, that the actors may modify their judgements through pressure to conform, or merely to please the analyst, rather than because they really do accept the requirements for consistency inherent in the decision aid. The precautions which were incorporated into the research method to cope with these problems are discussed in section 7.4.2.

7.2 THE PROTOTYPICAL CASE VIGNETTES

Three prototypical case vignettes were produced for this study (see appendix 4). Two of them were adapted from case vignettes used in the study described in chapter 6 and a further new case was produced. Like those in chapter 6, the vignettes were summaries of the information relevant to technical design, and they were devised to invite between them the potential use of a range of air conditioning system types for open plan/modular office spaces.

7.2.1 Production of the Case Vignettes

The case vignettes were once again produced drawing upon the general knowledge of what constitutes the 'typical' form of design description for office buildings at the beginning of the feasibility stage gained from the preliminary study of chapter 3. This time, however, the vignettes were further tailored to ensure that comfort air conditioning was a reasonable requirement in every case using the information on how engineers determine fundamental HVAC requirements gleaned from the second study of chapter 6. The engineers from the first study who reviewed the vignettes for the second study also assisted by reviewing the vignettes for this third study. They were asked again to comment on realism and completeness, so far as this is possible in a vignette, and to ensure that a wide range of air conditioning system types could be used. They were not, however, asked to state any preferences.

7.2.2 The Artificiality of the Vignettes

The vignettes for this study, like those used in the previous study reported in chapter 6, concentrated on the technical perspective of the decision situation and ignored complicating factors such as the political background to the case. While this is obviously an artificial situation, the success of the technique in the previous study indicated that the artificiality would not be of such effect as to render the results of the study too unrepresentative of a real situation.

Unlike the previous study, however, the vignettes were used in conjunction with an imposed decision making technique which was alien to the engineer subjects. The effect which this would have on the assessed validity of the MAVT technique was unknown, except that it is likely that the attendant simplicity will favour the technique appearing valid. Hence, should it be found that MAVT is *not* a valid technique when applied in this study context, it is unlikely that it could be valid if applied in a real situation.

7.3 ELICITATION TECHNIQUES EMPLOYED

The use of best practice in elicitation techniques has two objectives. The first is to ensure that the MAVT technique itself is applied soundly, otherwise research into the validity of the technique would be open to criticism on the grounds that the technique under investigation was not applied properly in the first place. The second objective is to minimise confounding facilitator effects by using elicitation techniques which are as neutral as possible. These objectives are compatible. Best practice is also neutral practice, because the ethics of decision analysis require that the analyst does not allow his personal preferences and beliefs concerning the domain to influence the course of the analysis. As will be appreciated from the comments in section 7.1 above however, there is a limit to how successful it is possible to be in achieving these objectives.

This section gives a detailed account of the techniques which were used during this study following the step by step description of the weighted sum MAVT procedure in section 5.2.3.

7.3.1 Stage 1 - The Elicitation of Attributes

It was pointed out in section 5.1.1 that the structuring of the decision problem by selecting the attributes by which the options are to be evaluated is not given much attention in the literature. The structuring process is seen as more of an art than a science, and relies upon the creativity of the actor (and anyone assisting him), coupled with his particular view of the decision situation. In this respect then, many of the issues involved in the selection of criteria are extra-rational. However, notwithstanding this the end product of the structuring process has a tangible form which can, in principle, be well defined, and it was indicated that mainstream decision analysis holds that a valid set of attributes should be *complete*, *operational*, *non-redundant* and have a *range of convenience* which covers all of the options under consideration.

Almost all subject based studies in decision analysis of the 'experimental' type outlined in section 7.1.1 above begin by presenting the actors with a definition of the set of attributes to use in the evaluation. While this is convenient for the researcher, an attempt to do the same in a study designed specifically to assess the validity of a decision analysis technique (rather than merely produce insights to the differences between techniques) could pose a threat to the methodological validity of the study. There is evidence to suggest that imposing a set of attributes on a decision actor in the first instance may not allow his particular view of the decision situation to emerge (Wooler, 1984). Moreover, it would not be possible to ensure that a list of attributes compiled *a priory* (eg culled from the HVAC design literature) contained all those relevant to a particular decision situation. " MOEs [attributes] will vary from evaluation to evaluation. As Dockerty (1985) puts it 'MOEs can only exist within some problem context'. For this reason, prescriptive lists of MOEs are inappropriate. They may encourage the inexperienced evaluator to use them as a matter of convenience or because they simply imply a prescriptive authority" (Riedel and Pitz, 1986).

The problem of how a facilitator or decision aid 'front end' can help actors generate their own set of criteria while remaining as neutral as possible was addressed by Humphreys *et al* (1980) using *repertory grid*, a technique based on Personal Construct Theory, for the early, formative stages of the analysis. It would seem that repertory grid remains the best available method of attribute elicitation to date, as witnessed by the repeated reference to it in the field of knowledge elicitation generally (eg Hart, 1986; Shaw and Gaines, 1987; Firlej and Hellens, 1991).

Personal Construct Theory is in its entirety a complex psychological model of how individuals perceive the world around them, however on a practical level its application through repertory grid for the purposes of knowledge elicitation is more straightforward (in this sense it is analogous to the technique of teachback interviewing introduced in chapter 6). Kelly (1955) is the original work on Personal Construct Theory while Fransella and Bannister (1977) and Stewart and Stewart (1981) offer more specific guidance on the application of repertory grid.

In repertory grid, the *elements* to be evaluated are usually compared in triads, and the *constructs* (the ways in which the elements are perceived by the subject) are elicited by asking the subject in what way one of the elements is similar to another yet different from the third. As an example (which was also used as an introductory example in the study) a subject could be asked to consider three modes of transport - a car, a train and a donkey (Stewart and Stewart). When asked how the car is similar to the train yet different from the donkey, the subject might reply that the car and the train run on wheels but the donkey walks on legs, or that the car and the train are fast while the donkey is slow. If asked how the car and the donkey are similar but different to the train, the subject might reply that they can go anywhere while the train must stay on the track, or that they carry few passengers while the train carries many, and so on.

Hence, the elements car, train and donkey have elicited the bi-polar constructs <run on wheels - walk on legs>, <fast - slow>, <go anywhere - stay on track>, <few passengers - many passengers>. For the purposes of MAVT, the elements are the options to be evaluated and the constructs give rise to attributes provided that they are *value constructs*, ie that they differentiate between the options in a way which makes one preferred to another (Brownlow, 1989). So in the above example it might be said that <run on wheels - walk on legs> does not qualify as a value construct because the fact that the element has wheels or legs does not *per se* affect its value as a mode of transport. The other constructs however could well be value constructs (depending upon the decision situation) and so would give rise to potential attributes of, say, speed, flexibility and passenger capacity. Whether these attributes would finally be chosen for use in the MAVT model, however, would depend upon their satisfying the standards of validity in section 5.1.1.

Each engineer who participates in the study will define an attribute set which is to some extent unique to him and possibly to each case vignette. It would be possible to continue the study using a unique set for each subject and vignette, but as will be seen, this would make comparison of their MAVT models extremely difficult at a later stage. Hence, it was decided to adopt the practice common in knowledge elicitation of combining the attribute sets into one consensual set at this point, provided this could be achieved without profoundly altering any of the individual sets. This was accomplished by using the non-parametric factoring procedure (see Coshall, 1991 for a readable explanation of this simple and powerful technique) to identify constructs which represented the same or similar concepts *within* each grid (and so might indicate hierarchical relationships and/or double counting) and then *between* each grid (and so might indicate a shared perception of the domain). The list of candidate attributes suggested by this consensual set was then presented to the individual engineers for refinement through more conventional discussion.

In the event, it was found that there was good agreement among the engineers on attribute definitions, and this part of the research method did not present serious difficulties. This is not surprising given that the decision making task study of chapter 6 indicated that the domain is well structured, with the engineers sharing an essentially similar form of decision strategy. Moreover, the list of attributes compiled informally in table 6.1 is a further hint that there is a high degree of agreement on problem structure. The application of repertory grid and specific results of this stage in the process are presented in section 8.2.

7.3.2 Stage 2 - Specifying Individual Attribute Values

The second stage in the elicitation process is to assign a value v_{ij} for the performance of each option i on each of the individual attributes j. The development of numerical scales which reflect a subject's perception of an object has been the topic of much research both in the fields of decision analysis and psychophysics. Some caution is required when developing value scales because it has been shown that subjects' responses can be distorted by several measurement biases. There is not the space here to consider these biases in detail (see Poulton, 1979 for a summary), instead, this section will concentrate on describing the elicitation methods used to overcome these biases.

It is generally accepted in mainstream decision analysis and in psychophysics that reasonable numbers can be elicited using an *anchored rating scale*; that is to say an interval scale which has a clearly defined upper and lower limit, and on which the value of objects is rated by comparison with those end anchors (Anderson and Zalinski, 1988; von Winterfeldt and Edwards, 1987; Watson and Buede, 1986). In the case of intangible attributes, this rating must be done directly to the numerical value scale, while for incommensurable attributes with a natural underlying measure x_j an attribute value function $v_j(x_j)$ is specified showing graphically how the value v varies with the underlying measure x.

Intangible Attributes

Suppose an actor who is choosing a motorcar from among a range of four options, A, B, C and D, has the attribute 'comfort' in his attribute set, and that no satisfactory measure of comfort could be found, making it an intangible. Further suppose that the rank order of the cars as assessed in comfort terms is A, B, C, D, with A being best and D worst.

The first requirement is to fix the lower and upper end anchors of the attribute value scale. Obviously these must be defined as at least as poor as D and at least as good as A respectively if all of the options are to be included in the evaluation. The practice of defining the end anchors as the worst and best levels of the attribute from among the range of options is theoretically correct, however it may be that at some later stage in the analysis a new option is introduced with comfort which is worse than D or better than A. Had the attribute value scale been initially defined with limits according to D and A, then the scale (and therefore the weight) would need to be re-calculated. Hence, the end points may be defined more generally as the *threshold* level (ie least acceptable level) and the *highest feasible* level or alternatively the *satiation* level (ie the level after which no increase in performance is seen as beneficial). So in this example it may be that the actor decides that he is not prepared to consider a car which is less comfortable than a basic Ford Escort, no matter what its compensating features on the other attributes, and that the highest feasible level of comfort in the circumstances corresponds to a top-of-the-range Vauxhall Cavalier, although these particular cars may not be actual options. These two ways of defining the range are illustrated in figure 7.1.

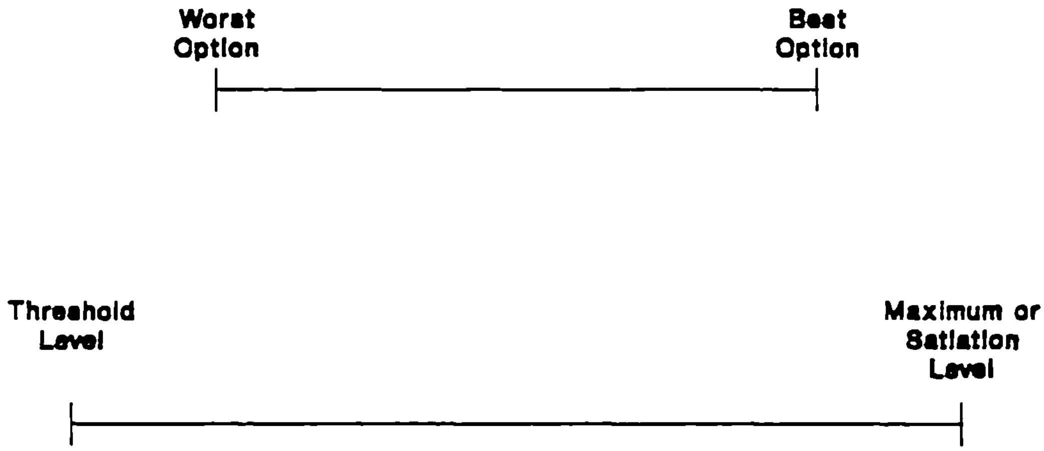


Figure 7.1 Two ways of defining the end anchors of an attribute value scale

The number of points in the value scale (eg 0-1 or 0-10 or 0-100) is theoretically irrelevant provided each attribute is given the same scale. The convention on the scale varies from one writer to another, but generally decision analysts have used 0-1 (eg Keeny and Raiffa, 1976; Watson and Buede, 1987) or 0-100 (eg von Winterfeldt and Edwards, 1986; Edwards and Newman, 1982). However, researchers in psychophysics have found that scales with fewer than 10 steps or more than 20 steps can induce certain response biases, and so normally 0-10 or 0-20 scales are recommended (Anderson, 1982).

Having defined the end anchors, the values of the options that lie between these end points are elicited. For intangible attributes this is simply done by asking the actor to award a score to each of the options, taking careful consideration of the end anchors. Again following the advice of the psychophysicists, this is best done using a graphic scale of the kind illustrated in figure 7.2 (Anderson, 1982) on which the actor can visually locate the various options with respect to the end anchors and each other.

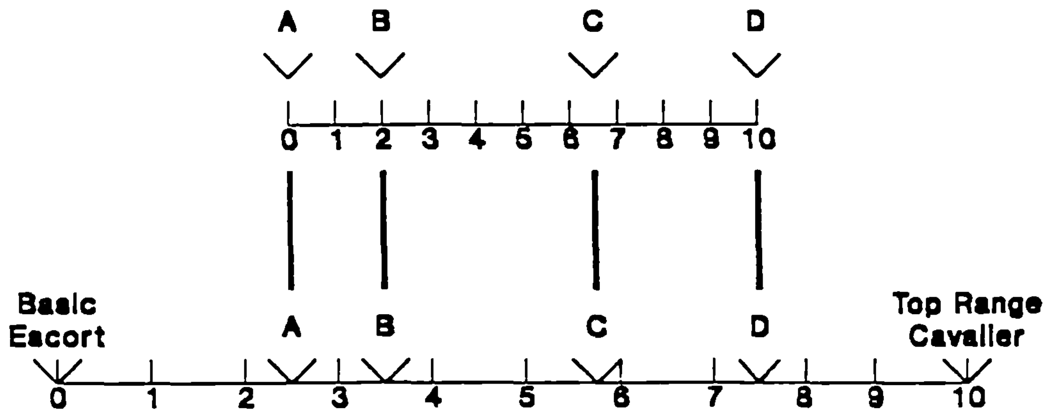


Figure 7.2 A graphic scale for scoring intangible attributes

Incommensurable Attributes

To extend the above example, suppose that the actor also has the attribute 'maximum speed' in his attribute set. This attribute has a natural measure in miles per hour, and so in this case the actor can define an attribute value function relating the underlying measure x to the value scale v. Further suppose that the four options have maximum speeds when fully laden as follows: Car A - 95 mph, Car B - 100 mph, Car C - 110 mph, Car D - 130 mph. The first requirement is to fix the lower and upper end anchors of the attribute value scale, and the considerations of which anchors to choose and the number of points in the value scale are the same here as for the intangible attributes described above.

Having defined the end anchors it is then necessary to determine how the value v of the underlying measure x varies between these end points. There are a number of ways in which this might be done (see Fishburn, 1967 and Kelly and Farquhar, 1974 for extensive reviews of possible methods) however the discussion here will be confined to the *bisection* method which has also been used in psychophysics and has been found to be less likely to induce bias. With the bisection method, the actor is asked to consider the measures which have been defined as the end anchors and then to find the point x_j which is half way *in value* between them. This point being established, the scale may then be further subdivided leading to refinement of the value function. It may be that the above actor has said he will not consider a car with a top speed of less than 90 mph, and that a top speed of more than 140 mph would be of no increase in value to him. In other words the threshold level is 90 mph and the value function becomes satiated at 140 mph. Next, it is determined that he considers the midpoint to be 105 mph - ie the improvement in top speed from 90 to 105 mph is worth the same as an increase from 105 to 140 mph. Following this, the midpoint of the difference between 90 and 105 mph is then determined, as is the midpoint between 105 and 140 mph. If these are found to be 95 mph and 120 mph respectively, then the value function is as plotted below in figure 7.3 and the scores for each car are: Car A - 2.5, Car B - 3.8, Car C - 6 and Car D - 9.

This is a classic shape of value function (referred to as *concave increasing*) often encountered in practical applications of decision analysis, resulting from the phenomenon economists refer to as diminishing returns to scale. The improvement in value v for an increase in x is rapid near the threshold point. As x increases however, the improvement in value for each unit increase in x reduces. From this graph the seemingly sensible inference may be drawn that the actor considers the improvement of 10 mph between 90 and 100 mph to be more valuable to him than the improvement of 10 mph between 130 mph and 140 mph. Examples of other forms of attribute value function also commonly found in decision analysis are illustrated in figure 7.4.

The use of the bisection method and the resulting value functions obtained in this study are described in section 8.3.

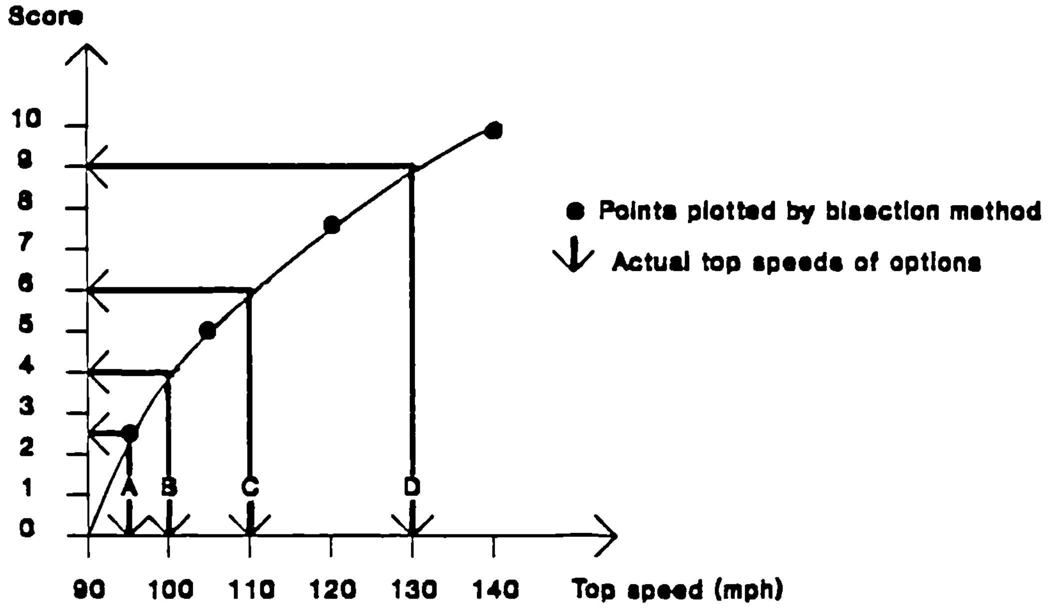


Figure 7.3 Illustration of a value function for 'top speed' plotted by the bisection method

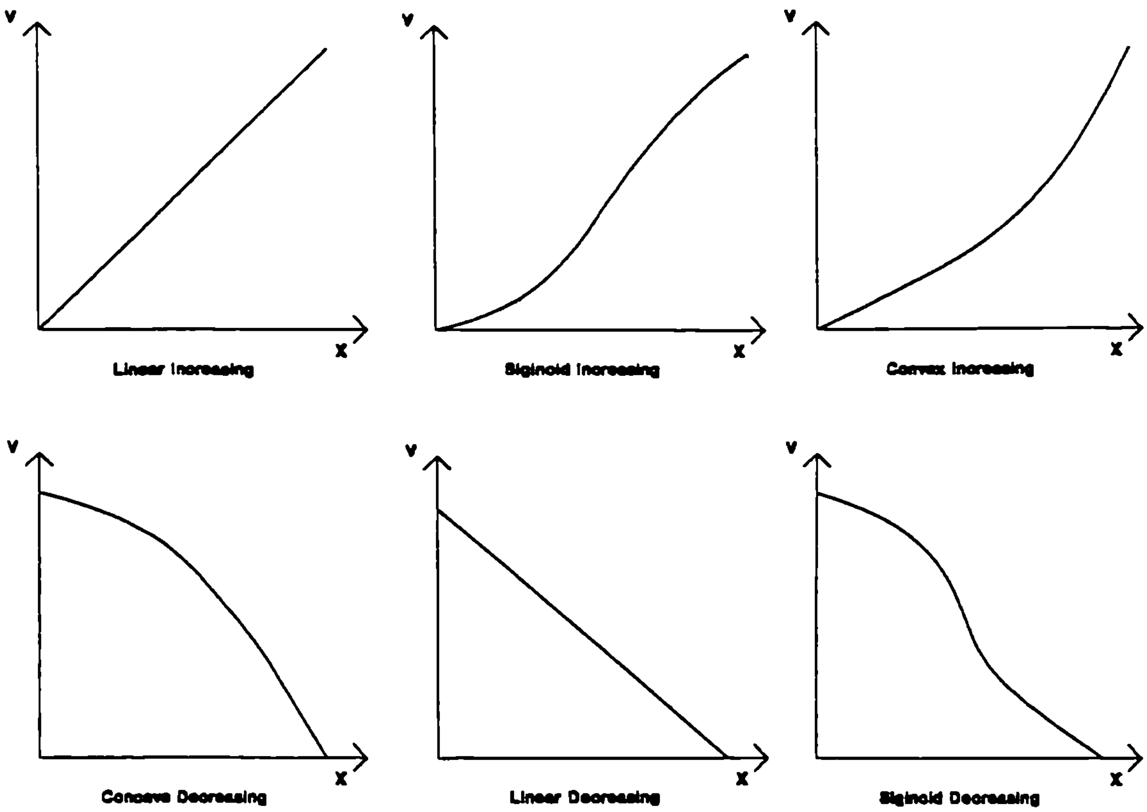


Figure 7.4 Examples of common forms of attribute value functions

7.3.3 Stage 3 - Verification of Additive Difference Independence

Once the values of each option on each criterion have been assessed, the criteria are checked for additive difference independence. To achieve this the analyst asks the actor to consider the difference in value of each criterion between two levels and ascertains that this perceived difference does not change as all the other attributes are swung from their worst to their best levels. For example, in a simple two attribute case an actor could be asked to consider his strength of preference for a car with a score of 7 for maximum speed and 0 for comfort over a car with a score of 3 for maximum speed and 0 for comfort. He would then be asked to consider the strength of preference for a car with a score of 7 for maximum speed and 10 for comfort over one with a score of 3 for maximum speed and 10 for comfort. If additive difference independence is to hold, the strength of preference should be the same.

If there are more than two attributes then each attribute must be independent of all the others for additive difference independence to hold. Formally, to carry out such an assessment for n attributes would require $n(n-1)/2$ comparisons which would be extremely tedious for a realistically large attribute set (n is generally between 10 and 20 in realistic applications). Hence in practice the analyst must apply judgement and concentrate the tests on potentially dependent attributes (see below).

If a violation of additive difference independence occurs, the analyst must examine the attribute definitions in an attempt to identify the source of the interaction between the attributes. Most writers on applied decision analysis suggest that interactions occur either because the attributes as defined overlap and contain facets of each other, or because an attribute which captures an interaction between other separate attributes has been omitted form the set (Watson and Buede, 1987; von Winterfeldt and Edwards, 1986; Keeney, 1981). This presumes, however, that the decision situation is not fundamentally unsuited to the application of MAVT as discussed in chapter 5. A more fundamental reason for independence being violated may be that the attribute ranges are too great, and a low score on one attribute cannot be compensated for by high scores on the other attributes. Should this occur it may be that an option has been included in the short list which has an inferior performance on one or more attributes which should have eliminated it at an earlier stage in the decision making process. The removal of such options (if this is justifiable) and the re-scaling of the offending attribute value scales may restore additive difference independence.

As an illustration of overlapping attributes, suppose that in the assessment of a motorcar an actor has defined two of the attributes as 'interior space' and 'comfort'. When independence of these attributes is tested it would be found that the more interior space a car has the more comfortable it is. The attributes overlap and so are not independent. In this case, the problem might be solved by re-defining 'comfort' more explicitly to exclude the common facet, perhaps as 'smoothness of ride' and 'support of seats'.

It should be noted that overlapping only relates to the dimensions on which the actor perceives value, and not to correlations between attributes which might exist in the domain. For example the attributes 'top speed' and 'acceleration' of a motor car can be considered independent even although the chances are that a car with a high top speed also has a high acceleration.

As an example of a missing interaction attribute, Keeney (1981) gives an illustration in which there are two equally weighted attributes. According to the decision actor, he would prefer an option which provides a balance of performance between both attributes rather than a high score on one and a low score on the other. On the face of it, this preference structure violates additive difference independence, however Keeney suggests that the problem has arisen because the attribute 'balance of performance' has been omitted form the attribute set.

The checks which were applied to verify additive difference independence in this study are detailed in section 8.4.

7.3.4 Stage 4 - Specifying Individual Attribute Weights

Most procedures for assessing attribute weights w_j begin by establishing a rank order on the attributes in terms of the swing in importance from worst to best (Watson and Buede, 1987; von Winterfeldt and Edwards, 1986; Keeney and Raiffa, 1976). The actor is asked to imagine that there is a hypothetical option which scores zero on all of the attributes, but he has the opportunity to improve its performance on *one* of the attributes so that it scores a maximum ten. Which attribute would he choose? After that one is improved, which would he improve second, then third and so on?

Once the rank order of the weights is established their relative magnitudes must be assessed. The literature discusses two or three ways in which the analyst might proceed, however the only truly theoretically valid method is to ask a series of trade-off questions. Since this research is concerned with the issue of validity, this is the method which was adopted here. Going back to the hypothetical options, the actor is asked to estimate how much performance on one attribute he would be prepared to sacrifice to gain a certain improvement in performance on another. For example suppose an actor is assessing a car on n attributes j, among which are 'top speed' ($j=1$), 'acceleration' ($j=2$) and 'smoothness of ride' ($j=3$). Furthermore, suppose that for the attribute ranges as defined by the actor it has been determined that $w_1 > w_2 > w_3 > \ldots > w_n$. The actor is then asked, starting with a top speed score of 10, how much in top speed would he be prepared to sacrifice to increase acceleration from its worst level (0) to its best level (10) assuming that all other attributes remain unchanged? He may answer that he would be prepared to loose 5 points on the top speed attribute to gain the 10 points on the acceleration attribute. This then implies that $w_1 = 2w_2$. Next the actor is asked for a trade-off between 'acceleration' and 'smoothness of ride' or between 'top speed' and 'smoothness of ride' to produce an equation involving w_3 and so on.

When this procedure has been repeated for each attribute there will be a set of n-1 equations relating the weights to each other. Finally, it is conventional to normalise the weights to sum to an arbitrary total such as 1 or 100. If 100 was chosen (as in the software package used in this study - see chapter 8) then the normalisation equation:

$$w_1 + w_2 + \ldots + w_n = 100$$

would be used to solve the set of n simultaneous linear equations.

In practice it is considered best to over specify the weights by asking more than n-1 questions, because trade-offs among the attributes go to the heart of making decisions with multiple criteria and so are difficult to estimate accurately. The additional trade-off questions can be used to find inconsistencies in the decision actor's judgements. A complete over specification would require $n(n\text{-}1)/2$ questions and for the same reasons as given in section 7.3.3 above would not be practical. However such checks can be concentrated on those attributes which have the highest weights (and therefore are more crucial to the validity of the model) and those for which the actor finds trade-offs particularly difficult to estimate.

Section 8.5 in the next chapter describes the application of this technique in the present study

7.3.5 Stage 5 - Analysis of the Total Weighted Scores

The final stage in the procedure is to select the option which is 'best' overall by analysing the total weighted scores. Formally, the option with the highest score is the preferred option, however it is often the practice in decision analysis to leave cost out of the value model constructed in stages 1-4 above and have the cost trade-off conducted as a separate exercise against the total weighted score of all the other performance related attributes. The reason for this is that most writers in applied decision analysis have observed that trade-offs involving cost (which may be capital or life-cycle) are highly sensitive and unstable (Watson and Buede, 1987; von Winterfeldt and Edwards, 1986; Edwards and Newman, 1982) and recent experimental work bears this out (Borcherding *et al*, 1991). Cost, when treated as an attribute in the normal sense, is invariably weighted very highly, and so when attempting to make trade-offs with other attributes which may have weights an order of magnitude lower, slight variations in judgement can produce very large changes in these other attribute weights.

As an example of such an analysis, suppose that an actor choosing a car has arrived at a total weighted score for all attributes except cost, and that the total score and cost (capital for the purposes of this illustration) of each car are as shown below in table 7.1. From this information it is possible to plot total weighted score against cost as in figure 7.5 to facilitate the final performance to cost trade-off.

Because the performance attributes are measured on an interval scale it is not possible to know in any objective sense which option is the best value for money. This decision, like all others in MAVT, depends on the judgement of the actor. Nevertheless, there are certain rational principles which can assist in this task by identifying what is termed the *efficient* options.

From the graph it is possible to see immediately that Car C is *dominated* by Car B because it costs more but has a lower total score, and Car F is likewise dominated by both Car D and Car E. These dominated options can therefore be eliminated from the assessment.

Next, note that if the actor believes that it is worth paying an extra £1,500 to gain an additional 0.7 value points (£2,128/point gain), then he will choose Car D over Car B, however such a judgement automatically implies that he will also prefer Car E over Car D as the additional 1.0 value points only costs £200 (£833/point gain). This means that, although it is not dominated, Car D is not a contender option either. Hence, the only efficient options are A, B, E and G and they form what is termed the *efficient frontier*, indicated by the lines connecting them in the figure.

OPTION	COST	TOTAL SCORE
Car A	8,800	4.5
Car B	9,000	5.5
Car C	9,500	5.0
Car D	10,200	6.2
Car E	10,700	7.2
Car F	11,500	6.0
Car G	13,000	8.0

Table 7.1 Total weighted score and costs for each option in the motorcar illustration

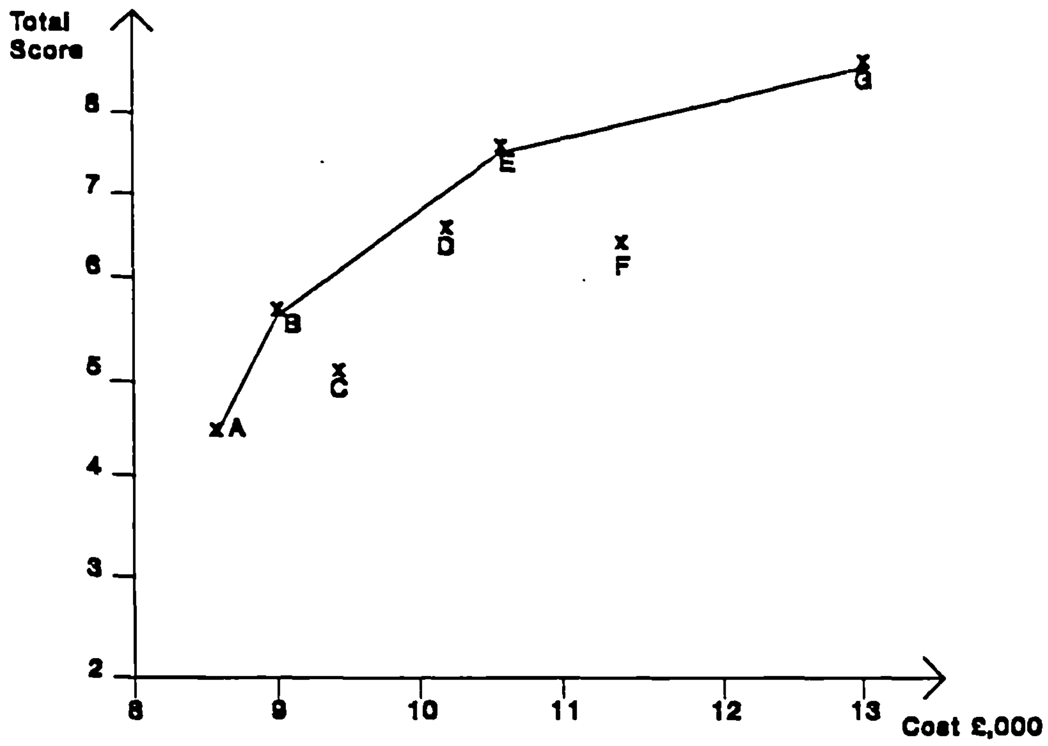


Figure 7.5 Plot of total weighted score against cost in the motorcar illustration

Having identified the efficient frontier, the choice of option depends upon the point at which the actor decides that the improvement in value between options no longer justifies the increase in cost. In this illustration it might be that the actor believes that it is worth paying an additional £1,700 to gain a further 1.7 value points (£1000/point gain) in which case he would prefer Car E to Car B (and therefore Car A). He may not believe that it is worth paying a further £2,300 to gain 0.8 value points (£2857/point gain) and therefore he would not choose Car G over Car E. Car E would therefore be judged the best option.

It should be noted that this form of analysis assumes that the value the actor attaches to money is linear across the range of costs under consideration - ie the £1,000 difference between £8,000 and £9,000 is worth the same to the actor as the £1,000 difference between £12,000 and £13,000. The literature generally assumes that this is the case, but there may be circumstances where this is not so, such as increasing reluctance to spend higher amounts (Keeney and Raiffa, 1976). If there are non-linearities in the value of money, then an attribute value function for money can be constructed and its value scores v_{ij} used in the above form of analysis instead of the raw cost figures.

The analysis of the total weighted scores in this study is reported in section 8.6.

7.4 TESTS FOR VALIDITY OF THE MAVT MODEL

In this study, the theoretical and normative validity of the MAVT models was tested in two ways. Essential tests were conducted as part of the model building process in keeping with the requirements inherent in MAVT described in section 7.3 above, and some additional tests were applied as part of the research study methodology.

7.4.1 Validity Tests forming part of Model Construction

In summary, the tests of theoretical and normative validity which were applied as part of the MAVT process are:

Theoretical Validity

1 A set of criteria can be specified which have an appropriate range of convenience.

2 The range of the natural measure x of a value function encompasses all of the relevant options.

3 The model can be constructed to satisfy additive difference independence.

4 Weights can be specified which are consistent with each other.

Normative Validity

1 The criterion set is complete (in terms of the view of the decision situation adopted).

2 The criteria are operational

3 The criterion set is non-redundant.

4 The value functions imply an intuitively correct variation of perceived value v with the natural measure x.

5 The implications contained in the model do not clash with intuitive judgements which the actor is confident in making.

7.4.2 Validity Tests forming part of the Research Methodology

The use of six subject engineers and three case vignettes allowed comparisons to be made between the judgements of each engineer from which the validity of MAVT could be inferred.

Inter-actor comparisons

It has been proposed that one criterion by which the validity of a decision analysis method might be tested is inter-actor convergence. This criterion assumes that evaluations made by fully knowledgable and rational people should agree. For example Newman (1975) suggested that for a decision analysis method to be valid it should produce greater convergence in the decisions of different expert actors than there is among their unaided decisions. Hobbs (1986) criticised this measure of validity on the grounds that by this standard, simply applying equal weights to all of the attributes would yield the most 'valid' MAVT models, and yet this is obviously nonsense.

In addition to this argument, it might be added that forced consistency rather than improved decision making could be responsible for the increased agreement. Since differences in opinion can exist between even expert actors, simple uniformity in judgement cannot be used as a validity criterion.

Notwithstanding this, it has been found in the field of expert knowledge elicitation that reasonable agreement can be expected in restricted domains. The more restricted the domain, the greater the agreement is likely to be. (Hart, 1986; Firlej and Hellens, 1991). Indeed, it is only because experts tend to agree in restricted domains that the development of expert systems is possible (Waterman, 1986). Hence, although it may be wrong to expect agreement between actors on the final decision of which air conditioning system type to use in certain circumstances (the wider domain), it is reasonable to expect some agreement on more restricted elements of the appraisal such as the form of value functions, the relative scores given to each option on each attribute, and the relative weights given to attributes. In this study, it was decided that while actual agreement on value functions, scores and weights would be unreasonable to expect, serious *disagreements* on these aspects may indicate a lack of validity in the model and warrant investigation. For example, if one actor's value function $v_j(x_j)$ indicated an increasing attribute value v with natural measure x while another actor's value function indicated the opposite, then this is more likely to be due to a misunderstanding during the model construction process than a genuine difference of opinion.

Intra-actor comparisons

Pitz *et al* (1980) suggested that the essence of validity in a decision model is the degree to which the model is sensitive to information about the decision situation (the case vignettes in an experimental setting or in a study such as this). If the model is valid, then it should exhibit superior responsiveness to unaided judgement. In the present study, it is possible to ask actors to make unaided judgements prior to conducting the MAVT analysis and then compare those judgements with the output of the MAVT model. It is possible that carry-over effects will distort the MAVT process; that is, having already made a judgement, the actor may subconsciously manipulate the MAVT process in an attempt to produce the same outcome. (Note that Pitz *et al* used a large factorial experimental design in an effort to control for carry-over effects and in the event found that no detectable effect existed. Other studies have also found no significant carry-over effects (Belton, 1986)). Although potential carry-over effects place a lower bound on the sensitivity of this test, it was nevertheless believed that it was essential to carry out such a test as it goes to the heart of the assessment of validity.

Chapter 8

A Study of the Validity of MAVT Applied to Air Conditioning System Appraisal

This chapter describes the implementation of the methods described in chapter 7. Six senior consulting engineers, who were not involved in the previous studies, participated in this study. The implementation of the methods required several interviews which were conducted separately at their consulting offices. To expedite certain stages in the construction and testing of the MAVT models, a lap-top computer running an interactive weighted scoring analysis programme called *Visual Interactive Sensitivity Analysis (VISA)* was used. This programme does not actually assist in the elicitation of attributes, or in the specification of raw weights and scores, however it does facilitate rapid calculation of normalised weights and of total weighted scores which would be laborious if attempted manually. This rapid calculation capability also facilitates sensitivity analysis of the total weighted scores to changes in weights or scores on a 'what if' basis. Data can be displayed as an *X-Y* plot, a feature which was adapted in this study to help with the final trade-off between total weighted score and cost. The VISA programme is described in detail in Belton and Vickers (1990).

8.1 ORIENTATION DISCUSSION AND PRELIMINARIES

Prior to embarking on the research interviews, some initial preparation was required. An informal orientation discussion similar to that described in section 6.1.1 was held in which it was explained that the research was aimed at assessing the 'suitability' (a term more meaningful to the engineers than 'validity' and sufficiently precise for a brief explanation) of the numerical decision aid MAVT to the task of appraising air conditioning system types. It was stressed that throughout the study the engineers would be recognised as experts. Even although the purpose of MAVT is to improve decision making, their judgements would form the basis by which the MAVT technique would be assessed, not the other way round. The engineers were informed as to the assumptions which were to be made regarding the domain view, and the level of detail at which the study would be conducted. It was explained that the processes of interest are those which guide the early stage selection of an air conditioning system type from among a range of options. Preliminary to the MAVT construction process, it had then to be decided which system options should be considered in the study, and how these would be ranked by the engineers using unaided judgement.

8.1.1 Determination of System Options for use in the Study

Following the orientation discussion, the engineers were presented with the three case vignettes shown in appendix 4 and asked to list and define the generic air conditioning system types of which they had detailed knowledge, and which they would consider using in each of the three cases. From their responses, a kernel list of the system types common to all the engineers was extracted and a working definition for each set down as shown in table 8.1

SYSTEM TYPE	ABBREVIATION	RELEVANT CASES
Traditional Variable Air Volume	**VAV**	**3**

DESCRIPTION

A central air handling plant with variable volume fan supplies a chilled mixture of fresh and recirculated air at conventional temperature to variable air volume control boxes installed in the ceiling void. Those boxes at the perimeter are fitted with hot water re-heater coils to offset winter heat losses. From the boxes the air is ducted down to ceiling mounted (fixed geometry) diffusers.

SYSTEM TYPE	ABBREVIATION	RELEVANT CASES
Fan Assisted Terminal Variable Air Volume	**FATVAV**	**2 & 3**

DESCRIPTION

A central air handling plant with variable volume fan supplies a chilled mixture of fresh and recirculated air at low temperature to series flow variable air volume fan boxes installed in the ceiling void. The fan boxes mix the low temperature primary air with locally recirculated air to produce mixed air at variable temperature and constant volume. Those fan boxes at the perimeter are fitted with hot water re-heater coils to offset winter heat losses. From the fan boxes the air is ducted down to ceiling mounted diffusers.

SYSTEM TYPE	ABBREVIATION	RELEVANT CASES
Four Pipe Fan Coil	**4PFC**	**1, 2 & 3**

DESCRIPTION

Fan coil units mounted in the ceiling void are supplied with chilled water from central chiller plant. Those fan coils at the perimeter are also fitted with separate hot water heating coils to offset winter heat losses. Separate air handling plant supplies fresh air for ventilation which is ducted to the fan coil units where it is mixed with locally recirculated air. From the fan coil units the air is ducted down to ceiling mounted diffusers.

SYSTEM TYPE	ABBREVIATION	RELEVANT CASES
Variable Refrigerant Volume	**VRV**	**1 & 2**

DESCRIPTION

Direct expansion (dx) fan coil units mounted in the ceiling void are connected in groups (usually of between four and eight) to an outdoor compressor/condenser unit. This unit and the entire* group of associated dx fan coils is capable of reversing the evaporator and condenser coil operation to pump heat to the perimeter. The dx fan coils handle locally recirculated air which is ducted down to ceiling mounted diffusers. Separate air handling units and ductwork supply fresh air for ventilation via independent ceiling mounted diffusers.

* Note that the system described has since been improved upon and it is now possible to achieve mixed heating and cooling from dx fan coils connected to the same outdoor unit. However at the time of this study, this was a new innovation and consequently not all of the engineers had a detailed knowledge of the characteristics of the new version.

SYSTEM TYPE	ABBREVIATION	RELEVANT CASES
Reversible Heat Pump	**RHP**	**1 & 2**

DESCRIPTION

Reversible heat pump units mounted in the ceiling void are interconnected by a two pipe water loop. Heat is extracted from or rejected to the water loop by the units while the water in the loop is held at a constant temperature by central boiler and heat rejection plant, depending upon whether there is a net gain or loss of heat in the building. The heat pump units handle locally recirculated air which is ducted down to ceiling mounted diffusers. Separate air handling units and ductwork supply fresh air for ventilation via independent ceiling mounted diffusers.

Table 8.1 Definitions of air conditioning system types

It was necessary to restrict the analysis to those system types for which all the engineers had detailed knowledge so that their respective MAVT models would be comparable, and so that the scores for each system type would be based on direct experience and not assumptions based on published information about the system types (which would have threatened the construct validity and reliability of the research). A negative consequence of this was that more unusual systems such as chilled ceiling and under floor VAV could not be included in the analysis. However, the systems considered do represent the range of types installed in the vast majority cases.

8.1.2 Unaided Judgements of System Types

In preparation for later comparisons between unaided judgement and those synthesised using MAVT, the engineers were asked to consider carefully the systems in table 8.1 and to rank them in order of preference as suitable air conditioning systems for each case vignette using only their intuition. It was stressed again that this preference order should be based on technical considerations only, and background political considerations should be ignored.

The engineers were given as much time as they required for this, as their deliberations took place between interviews. The engineers then wrote their preference rank orderings down on a card and placed it in an envelope. On collection at the next interview, the researcher was therefore unaware of the rankings as a precaution against any subconscious facilitator effects which may have resulted from this knowledge.

Comparison of these unaided judgements with those synthesised using MAVT are discussed in section 8.7.1.

8.2 ELICITATION OF ATTRIBUTES

This section details the application of the repertory grid technique described in section 7.3.1 followed by a further interview in which a consensual set of attributes was agreed upon.

8.2.1 The Repertory Grid Elicitation Process

The names of the system types in table 8.1 were written down on cards and used as elements in a repertory grid elicitation. The principles of construct elicitation were first illustrated using 'car', 'train' and 'donkey' as a triad of elements (see section 7.3.1). When the engineer understood the technique, triads were formed at random from the five air conditioning system types and the constructs so elicited were entered into a repertory grid.

By way of example, figure 8.1 presents an extract from a grid elicited form one engineer. The *elements* (ie the generic system types) appear at the top left hand side. The three systems used to elicit a construct are indicated by the circles in the boxes below the system headings. The ticks and crosses in these boxes indicate the ways in which two of the systems are seen as alike and yet different from the third. In the first row, the engineer was asked to state one way in which any two of the systems VAV, 4PFC and VRV were alike and yet different from the third in terms of their use in any of the case vignettes. As shown by the ticks and crosses in the first row, the engineer considered that VRV and 4PFC are alike in that they only require a narrow ceiling void (the positive pole of the construct) while VAV requires a deep void (the negative pole of the construct). One basic dimension of cognitive evaluation used by this engineer is thus identified. The engineer was then asked if the construct applies to any of the other systems and to which vignettes it applied (the construct may not apply to some systems or in some cases). As is shown in the first row by the other ticks and crosses, the engineer felt that the construct applies positively to RHP and negatively to FATVAV.

VAV	FATVAV	4PFC	VRV	RHP	CONSTRUCT		RELEVANT CASE VIGNETTES
					POSITIVE POLE (✓)	NEGATIVE POLE (X)	
⊗✗	✗	Ø✓	Ø✓	✓	Require only narrow ceiling void	require deep ceiling void	1, 2, 3
✓	✓	Ø✓	⊗✗	⊗✗	Require maintenance staff with normal training	Require specially trained maintenance staff	1, 2
⊗✗	⊗✗	Ø✓	✗	✓	Can create small control zones	Only larger control zones	1, 2, 3
⊗✗	✗	Ø✓	✓	Ø✓	Used everywhere	Only used on green field sites	1, 2, 3
✓	Ø✓	⊗✗	⊗✗	✗	Integral ventilation provision	Separate ventilation provision	1, 2, 3
Ø✓	Ø✓	✗	⊗✗	✗	Good air movement	Stagnation or draught prone	1, 2, 3
✗	⊗✗	⊗✗	✗	Ø✓	Take advantage of building orientation	Little advantage gained from orientation	1

Figure 8.1 Example extract from a repertory grid

He also believed that this construct was a valid consideration in all three case vignettes. Six other elicited constructs are also illustrated in this extract.

When the engineer could no longer think of any new constructs (usually after 30 or 40) the interview was terminated. The grids thus obtained from the six engineers were then analysed using the non-parametric factoring technique referred to in section 7.3.1. From this analysis, a trial list of attributes suggested by the core constructs was drawn up together with a list of candidate attributes suggested by the other constructs which had been rejected due to validity considerations such as double counting. These lists were then presented to the engineers for discussion in a third round of interviews, where agreement was sought for accepting or rejecting certain attributes, and on operational measures for the selected attributes.

As an illustration of how this process of analysis followed by discussion helped produce a valid set of attributes, it can be observed from figure 8.1 that the construct <can only use in green field sites - can use anywhere> mentioned by one engineer is totally positively correlated with the construct <deep ceiling void - narrow ceiling void>. It was also found that this construct was highly positively correlated with other space related constructs such as <large plant room area - small plant room area> in the same grid (not shown in the extract). The grid of two other engineers showed similar correlations between space related constructs and their construct <use in any building - use only in large building>. It was deduced from the grid analysis and following discussions that the underlying principle was some systems were less suited to buildings where space was at a premium than others. It so happened that those with a restriction on space were usually smaller and constructed in city centre locations while those with fewer space restrictions tended to be larger and on green field sites. Hence an air conditioning system's ability to service a building of a particular size or in certain locations are not attributes in themselves but merely *implications* of attributes (Fransella and Bannister, 1977).

	Attributes	Definition	Measure
1	Plant Room/Riser Area	Relative floor area occupied by air conditioning plant rooms and vertical risers for ducts, pipes etc	Percentage GFA
2	Ceiling Zone Depth	Depth of clear ceiling space required for horizontal runs of ducts, pipes, etc in addition to the depth required for suspended ceiling grids, lights etc	Depth in mm
3	Maintenance Nuisance	The amount of maintenance activity confined to plant rooms versus that which is intrusive and disruptive to the office space	Percentage of activity confined to plant rooms
4	Maintenance Specialisation	The degree of specialist knowledge required on the part of the maintenance staff	Intangible
5	Load variations handled	Variation in cooling loads from zone to zone which the system can cope with adequately while maintaining temperature control	Ratio of variation highest to lowest load
6	Ventilation Rates	Nominal ventilation rate for the system	Air change per hour
7	Air distribution	Stagnation or draught proneness of a system indicated by the difference from ideal velocity of around $0.2ms^{-1}$ in the occupied zone (i.e. up to 1.8m from floor) which may be experienced in some areas	Variaiton above and below $0.2ms^{-1}$ in ms^{-1}
8	Maximum Humidity	The maximum humidity likely in the conditioned space	% relative humidity
9	Noise Production	Noise rating achievable after a few years running	NR curves
10	Control Zone Size	Typical area of each independent control zone	Area m^2
11	Load Increase Contingency	Long term increase in cooling load in certain zones which could be allowed for realistically in the design without excessive extra cost	% over design cooling load
12	Reduced Load Performance	Long term reduced load in certain zones which could be accommodated without excessive loss of system efficiency	% of design load

Table 8.2 The final attribute list

ATTRIBUTE	REASON FOR REJECTION
Running costs	Costs considered separately
Capital costs	
Maintenance costs	
Energy costs	
Space requirements	Explicit in 1 & 2*
Building size	Implied by 1 & 2
Temperature control	Explicit in 5 & 10
Humidity control	8, otherwise restricted range of convenience
Centralised plant	Implied by 3
Short term flexibility	Explicit in 5 & 10
Long term flexibility	Explicit in 11 & 12
Air quality	7, otherwise non-discriminating
Environmental friendliness	Implied by running cost and refrigeration risks
Floor area	Explicit in 1
Services zone	Explicit in 2
Maintenance	Explicit in 3, 4 and maintenance cost
Efficiency	Reflected in energy costs
Building orientations	Reflected in energy costs
Comfort levels	Explicit in 5, 6, 7, 8, 9 & 11
Equipment durability	Reflected in maintenance costs
Local control	Explicit in 10
Risk of water damage	Restricted range of convenience
Risk of refrigerant leakage	Restricted range of convenience

* numbers indicate attributes on the list in Table 8.2

Table 8.3 The list of rejected attributes

The result of this analysis and discussion process was that the list of attributes presented in table 8.2 was agreed upon. Table 8.3 shows the list of rejected attributes with an indication of the reason for rejection.

8.2.2 Validity of the Attribute Set

The attribute set in table 8.2 was reviewed to determine the extent to which it satisfied the requirements for validity discussed in section 5.1.1, namely that the set was *complete, operational, non-redundant* and that the attributes had a *range of convenience* which covered all of the options. It was believed that while the attributes were operational and non-redundant, the set was not complete because some of the candidate attributes in table 8.3 had to be rejected due to a restricted range of convenience. The candidate attribute 'risk of refrigerant leakage' was only felt to apply in any meaningful way to VRV, while the candidate attribute 'risk of water damage' was not believed to apply to VRV. 'Humidity control' presented a problem because although all systems could de-humidify to some extent, not all could humidify. Hence, while 'maximum humidity' could be used as an attribute, the more general 'humidity control' could not.

Hence it had to be conceded even at this early stage that it would not be possible to develop an MAVT model which was complete in the decision analysis sense. Consequently, some design considerations would have to remain external to the model.

In addition to the above, it seemed to the researcher on purely *a priory* grounds that some factors were missing from the analysis because constructs which would have given rise to them never appeared during the repertory grid elicitation process. Principally, there were no attributes which related to the architectural/aesthetic sympathy of the air conditioning systems, or to partition integration or air filtration effectiveness. These points were raised with the engineers during the discussions about the attributes, and it was found that while these might be valid attributes in some instances, they did not differentiate between the systems being considered in this study. Because they did not differentiate, they could not emerge as constructs in the triadic comparison procedure.

8.3 SPECIFICATION OF INDIVIDUAL ATTRIBUTE VALUES

Now that there was an agreed list of attributes, the individual attribute value scores for each system type could be specified. A further interview was held during which the procedures described in section 7.3.2 were employed. Value functions were constructed for all of the attributes except 'maintenance specialisation', which was treated as an intangible because no natural measure could be found.

8.3.1 Construction of the Attribute Value Functions

The attribute value functions were constructed using the bisection method. The upper end anchor was defined as the best feasible or satiation level on the attribute, while the lower end anchor was defined as the worst acceptable level. For some attributes, the value function was not the same for all three case vignettes, in which instance more than one value function was constructed.

During the construction process it was found that there was some initial resistance to expressing opinions in numbers which might imply an unwarranted precision. This is not uncommon in those who encounter such techniques for the first time (Watson and Buede, 1987; von Winterfeldt and Edwards, 1986) and the engineers were assured that there would be opportunities to test the consequences of uncertainties in the numbers at a later stage in the process (this is the prime feature of the VISA programme).

The engineers were given the opportunity to smooth out irrational kinks in the curves implied by the values expressed in the bisection method. It was found that such kinks occurred mainly in the initial stages of the interview and were more severe with those attributes which had a small natural range (eg ceiling zone depth). This was because the engineers tended to think in terms of increments which were relatively large compared with the whole range (such as increments of 25mm in a range of only 100mm).

However, with increasing experience with the bisection method, this tendency diminished. The functions for each engineer, attribute and (where necessary) case vignette are shown in appendix 5.

The individual attribute scores, rounded to the nearest whole number, derived from the value functions or awarded directly for attribute 4 'maintenance specialisation' can be found in the score columns of the tables in appendix 6.

8.3.2 Validity of the Attribute Value Functions

The theoretical and normative tests of validity relevant to this stage in the MAVT process require that the range of the natural measure x covers all the relevant options and that the value functions imply an intuitively correct variation of value v with natural measure x.

On inspection of the functions it was found that in all instances the range did cover all of the options in table 8.2 which were relevant to each case vignette.

The form of the value functions, once any initial kinks had been smoothed out, all seemed to represent a sensible variation of value v with natural measure x. While there was evidently some difference of opinion among the six engineers both about the range and curvature of the functions, no curve is a gross variance with the others or with common sense (eg increasing with x while others decrease or implying a higher score for an obviously inferior value of measure x).

8.4 VERIFICATION OF ADDITIVE DIFFERENCE INDEPENDENCE

As indicated in section 7.3.3 it is not practical to check every combination of attributes, however inspection of the attribute value functions can indicate possible problem areas. In particular, the analyst must be vigilant for overlapping attributes or attributes with excessive ranges on the natural measure x, both of which can lead to violations of additive difference independence.

8.4.1 The Search for Overlapping Attributes

An inspection of the list of attributes in table 8.2, the value functions in appendix 5, and the scores awarded by the engineers for the various system options suggested some possible overlapping attributes which were investigated. These were namely:

Plant room/riser area	and	*Ceiling zone depth*
Load variation handled	and	*Reduced load performance*
Ventilation rates	and	*Air distribution*
Ventilation rates	and	*Maximum humidity*

It was found that all of the combinations with the exception of *load variation handled* and *reduced load performance* were believed by all the engineers to be correlated in the domain, and not actually overlapping in a value sense. For example, it so happens that all-air systems such as VAV and FATVAV require more plant room and riser area and also greater ceiling zone depth. However when the questioning procedure described in section 7.3.3 was used, all of the engineers agreed that the value of the scores on one attribute were unaffected by changes in the scores on the other. It could therefore be concluded that these attributes were distinct aspects of value in the perception of the engineers.

No such unanimous agreement was found for *load variation handled* and *reduced load performance* however. Four engineers adhered to their original belief (when the attribute list was first reviewed) that these were distinct aspects of value, but the remaining two engineers felt that on further reflection these attributes might have a degree of commonality. As this doubt took the form of an unease rather than a clearly identifiable problem, it was agreed to note the reservation, but continue with the MAVT process using the attributes.

8.4.2 The Search for Excessive Attribute Ranges

Value functions with relatively large and relatively small ranges of the natural measure were selected and set against each other to test for possible excessive ranges which would preclude free trade-offs. For the purpose of the exercise, the size of the range was taken to be the ratio of the highest to lowest point on the natural scale (except for the attribute *air distribution* for which such a ratio is meaningless). The attributes with the three largest and three smallest ranges were thus identified as:

Large: *Load variations handled*
Ventilation rates
Load increase contingency

Small: *Ceiling zone depth*
Maximum humidity
Noise production

Attributes from each group were paired at random and the questioning procedure described in section 7.3.3 was used to probe for any changes in the score for one attribute due to a change in the score on the other. It was found that the scores were unaffected in this analysis. It could therefore be concluded that none of the attribute ranges was so large as to preclude valid trade-offs.

8.4.3 Conclusion of the Verification Process

Because two of the engineers had reservations about the overlap between the attributes *load variations handled* and *reduced load performance*, it is not possible to claim that additive difference independence is completely satisfied. However with the application of prescriptive methods some discretion is necessary, and the severity of the violation must be considered. Given that the two engineers merely expressed unease, but could not specifically articulate the nature of the overlap, it was considered that any potential violation is likely to be slight, and not affect the validity of the MAVT model to any serious extent. Nevertheless, caution in dealing with these attributes at subsequent stages of the MAVT process was indicated.

8.5 SPECIFYING INDIVIDUAL ATTRIBUTE WEIGHTS

This section describes how the attribute weights were specified using a single pass through the attributes using trade-off questions followed by a verification procedure. As section 7.3.4 explained, it is not practical to verify consistency with every combination of attributes, and so a pragmatic procedure was adopted whereby only the higher weighted attributes or those for which trade-offs were particularly difficult were subjected to verification.

8.5.1 First Pass at Specifying the Weights

For each case vignette in turn, the weight specification procedure of section 7.3.4 was applied. Great care was taken to emphasise the difference between 'importance' weight and 'trade-off' weight, as it was found that the notion of importance weight was firmly instilled in those who had encountered weighted scoring type analyses before. With careful initial explanation and occasional reminders it appeared that this bias could be overcome.

The engineers were first asked to rank order the weights, and once ordered, the relative magnitudes were assessed using trade-off questions between pairs of attributes for which the engineer felt most comfortable making comparisons. For example, no engineer suggested trading off *plant room/riser area* with *noise production* in the first pass, however most were comfortable trading off *plant room/riser area* with *ceiling zone depth* because the notion of sacrificing one form of space for another was more familiar. As individual weight relationships were inferred, they were entered into the visa programme where the complete set was finally normalised to sum to 100.

8.5.2 Verification of the Weights

It was obvious during the interview that assessing the weights was more difficult than specifying value functions. Those attributes which had the highest weights and those for which the engineers found trade-offs particularly difficult to make were noted and checks on their magnitudes were made by asking different trade-off questions. Inconsistencies were usually found during this procedure. The inconsistency was brought to the attention of the engineer and the weights were adjusted to reduce the inconsistency until it was judged to be within acceptable limits.

The notion of what constitutes an acceptable limit to inconsistency is not discussed in the decision analysis literature. Many writers suggest accepting the weights obtained in the first pass and then testing the robustness of the first ranked option to changes in weights at a later stage in the MAVT process (von Winterfeldt and Edwards, 1986; Belton, 1986; Belton 1985). This has the advantage of concentrating verification on those weights most likely to change the outcome of the evaluation, however these recommendations relate mainly to field applications of MAVT. This study is concerned with a more rigorous assessment of method validity, and so such an approach would leave too many untested implications (see section 4.3.3). An extension of this form of robustness test or *sensitivity analysis* was used in this study at a later stage, but at this point in the process it was decided to adopt a tolerance criterion for acceptable limits of inconsistency in the weights which would hopefully promote the more rapid convergence to valid weights.

The criterion adopted at the outset of the study was that the variation in the highest weight should not fluctuate under different trade-off questions by more than the magnitude of the lowest weight. That is to say, it should not be considered acceptable that the uncertainty (due to weight) in the weighted score of the most highly weighted attribute exceeds the entire swing in weighted score from worst to best on the lowest weighted attribute. The proportional variation thus established would then apply to all of the weights.

In the event, it was found that the weights of the highest and lowest weighted attributes were quite extreme, typically in the ratio of 30:1 and in the most extreme case 47:1. It proved impossible to achieve consistency to a tolerance of around 3.3%, and so the arbitrary tolerance of 15% was used. In effect, this variation in weights suggests that many of the lower weighted attributes are better viewed as *constraints*, whereby little value is derived from improved performance beyond the threshold level (Belton, 1985).

Problems of inconsistency were found to be worst when attributes with high weights were traded off against attributes with low weights, as slight variations in judgement could translate into large differences in the assessed weights. This situation was therefore avoided if at all possible.

The final set of weights specified by each engineer and rounded to the nearest whole number are shown in the weight columns of the tables in appendix 6. As comparison of the tables indicates, there is some disagreement among the engineers on attribute weights, however no set is at such variance with the others as to indicate an unfeasibly large divergence of opinion. For example there is broad agreement on the top third, middle third and bottom third weights as ranked by magnitude across the engineers, although the individual rankings and magnitudes vary.

8.6 ANALYSIS OF TOTAL WEIGHTED SCORES

The scores derived for the attributes in section 8.3 were multiplied by the weights specified in section 8.5 and these weighted scores were then summed to produce a total weighted score for each air conditioning system option. Following the procedure described in section 7.3.5, these total weighted scores were then compared with cost so that a final performance versus cost trade-off could be made and the 'best' overall option identified.

8.6.1 Option Ranks Resulting from the MAVT Models

The VISA computer programme rapidly calculated the total weighted scores and these scores were then displayed graphically against cost as an *X-Y* plot. (Note that VISA has no automatic facility for displaying total weighted score against an additional attribute such as cost. However, it was a simple matter to perform a separate evaluation in which total weighted score and cost were treated as two new attributes.) The system costs could be defined as either capital or life-cycle costs and so both were investigated. The estimated costs for each system/vignette combination were provided by two large firms of cost consultants with particular expertise in building services, and are based on prices as of November 1992 (this phase of the research was conducted between December 1992 and February 1993).

The graphs of total weighted score versus cost for each engineer are shown in figure 8.2. The life cycle graphs shown are based on a time horizon of eight years and a discount rate of five percent, these representing the upper and lower bounds respectively which the engineers believed that clients would be prepared to consider. It was generally thought that capital cost was viewed, even in owner occupation projects, as more significant than life-cycle cost. Hence, only a very powerful argument would persuade a client to spend more capital in the interests of reduced running costs.

Using the rational principles of dominance, efficiency and finally cost per point increase in score, the engineers arrived at the decisions indicated on the graphs by an asterisk. As it transpired, most of the decisions were insensitive to discount rate or time horizon as can be seen from the graphs. On the two occasions where consideration of life-cycle cost did indicate a change in preference to that indicated by consideration of capital cost alone (case vignette 2, engineer 2 and case vignette 3, engineer 5), in neither case was the difference significant enough to warrant altering the decision based on capital cost.

In a real decision situation, the first ranked option is all that matters, the rank ordering of the others being academic. However, the rank order of the other options is important in a study such as this because this information relates to the other implications of the MAVT model which are evidence of its validity. In addition to identifying the most preferred options therefore, a complete preference rank ordering was established using a process of elimination. Once the first ranked option had been identified, it was eliminated from the analysis and the principles of dominance, efficiency and cost per point were re-applied to find the second ranked option and so on. The rank ordering so established for each engineer is summarised in table 8.4.

8.6.2 Sensitivity Analysis

The robustness of the rankings obtained using MAVT were further subjected to a sensitivity analysis. The objective of such an analysis is to determine how sensitive the outcomes are to changes in weights or scores about which there will inevitably be some uncertainty. It may be that weights and scores may be varied to the upper and lower bounds of the uncertainty without affecting the first ranked option, in which case the actor may take some comfort from the thought that reasonable uncertainties over scores and weights are not significant. However it may be that changes in rank order do occur due to changes in weights or scores, which indicates caution when interpreting the results of the analysis.

The VISA programme provides a facility for assessing the sensitivity of the outcomes to weights and scores by allowing the values to be changed interactively and displaying instantly updated results on the screen. When altering a weight, the total is automatically re-normalised by reducing or increasing the other weights proportionately (Belton and Vickers, 1990). It quickly became obvious during the research that no reasonable changes in weights or scores would alter the seemingly obvious decisions such

Total weighted score against capital cost

Total weighted score against NPV of capital and running costs over eight years

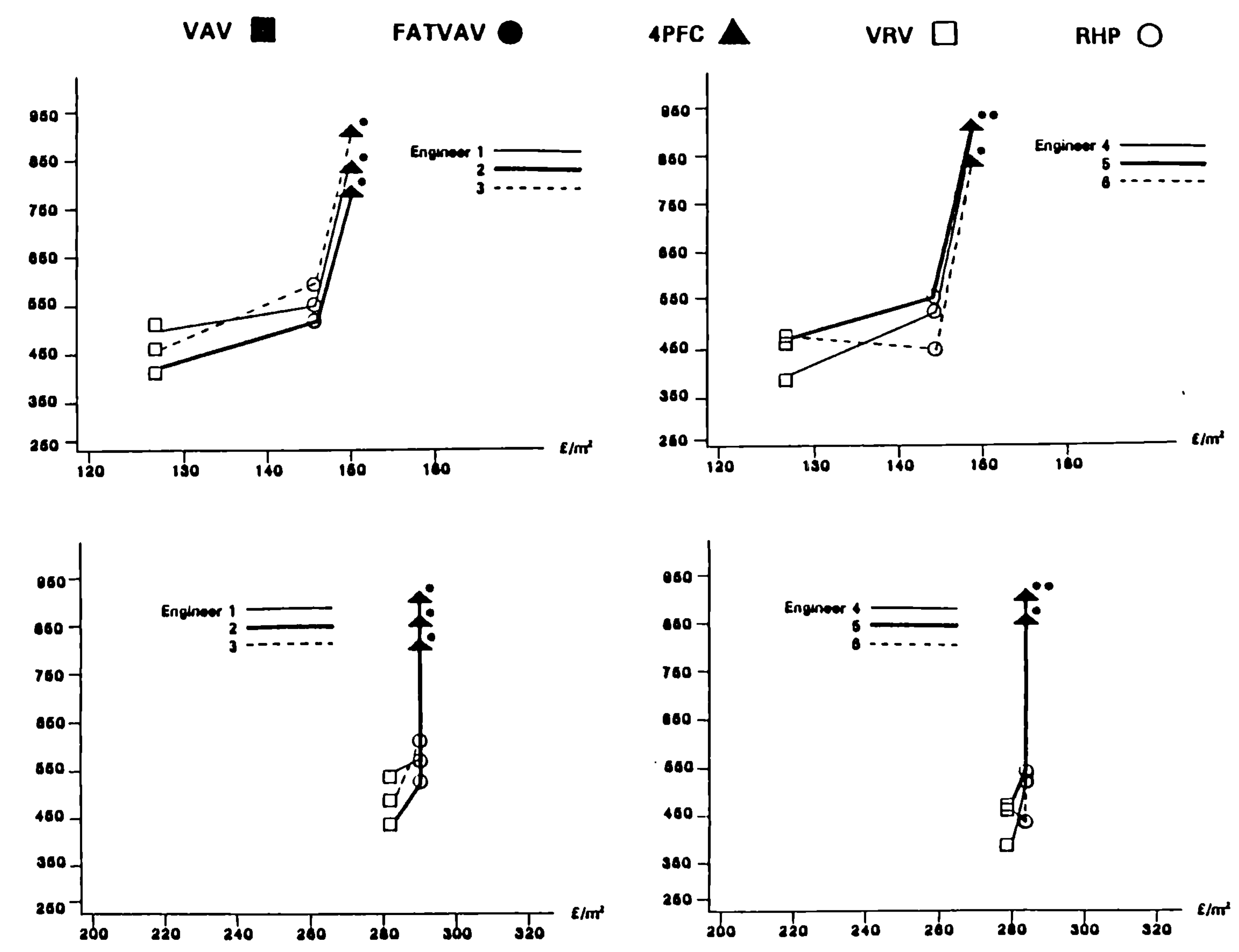


Figure 8.2 (a) X-Y plot of total weighted score against costs for case vignette 1

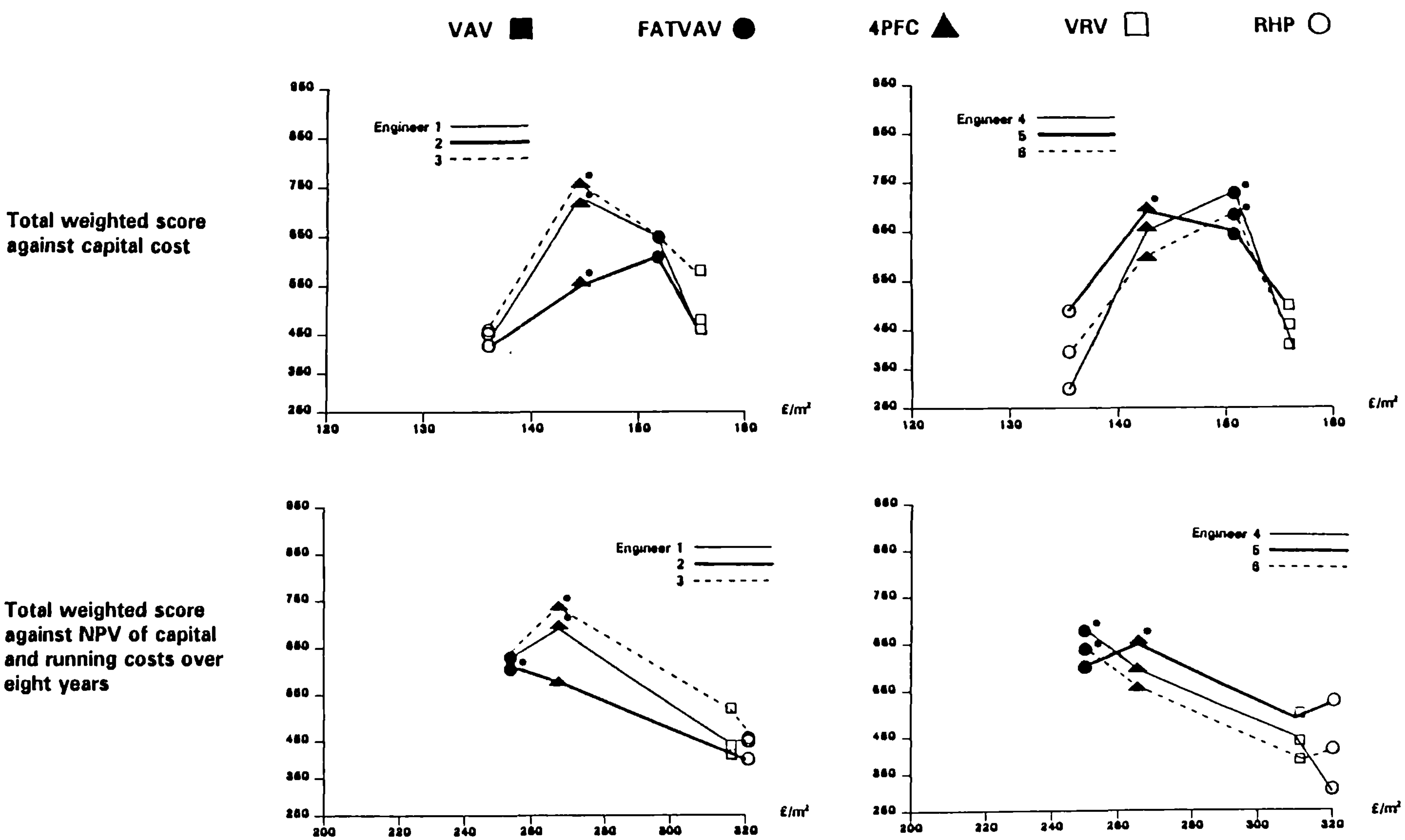


Figure 8.2 (b) X-Y plot of total weighted score against costs for case vignette 2

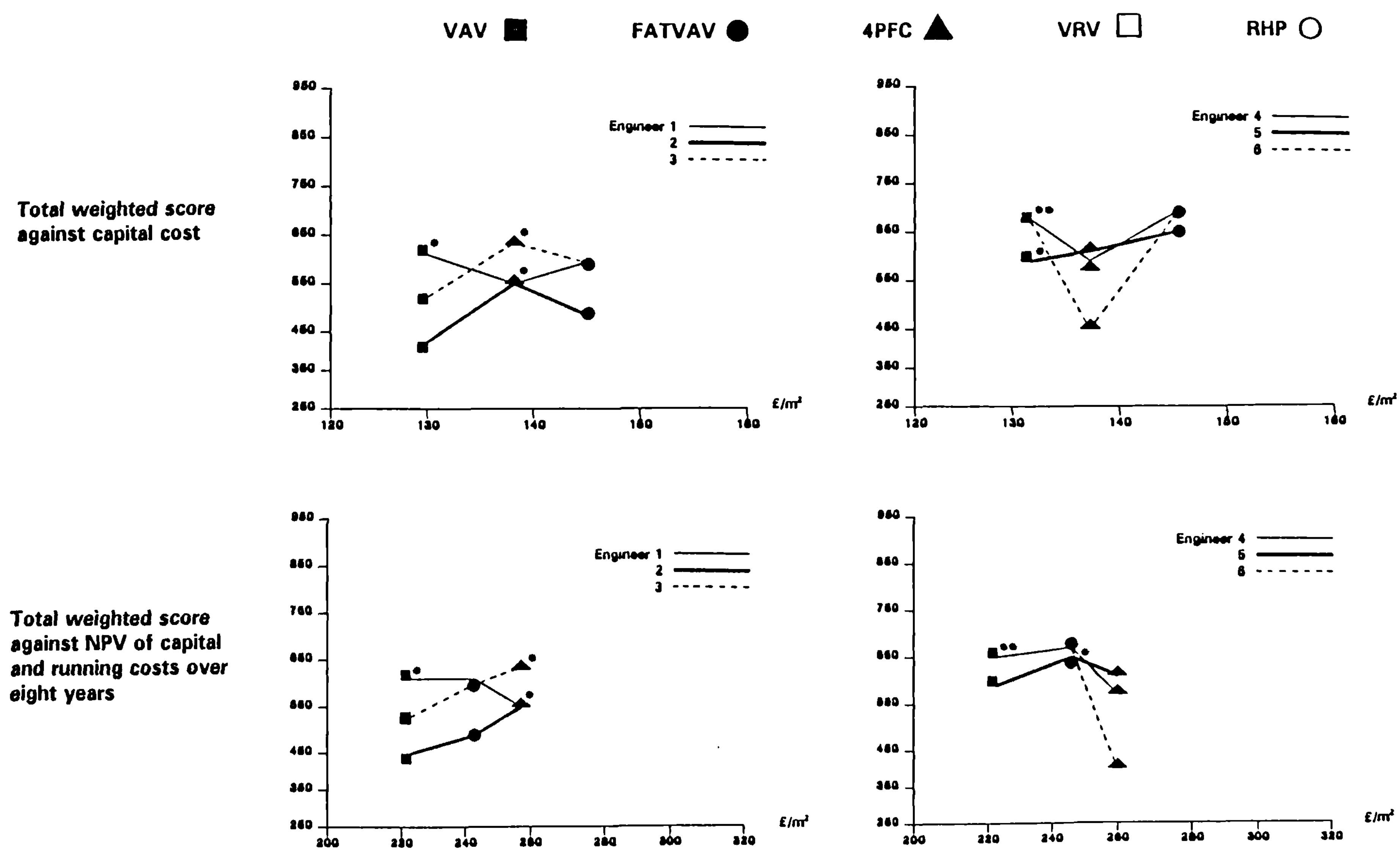


Figure 8.2 (c) X-Y plot of total weighted score against costs for case vignette 3

Engineer	Rank	Case Vignette 1	Case Vignetee 2	Case Vignette 3
1	1	4PFC	4PFC	VAV
	2	VRV	FATVAV	FATVAV
	3	RHP	RHP	4PFC
	4	-	VRV	-
2	1	4PFC	4PFC	4PFC
	2	RHP	FATVAV	FATVAV
	3	VRV	RHP	VAV
	4	-	VRV	-
3	1	4PFC	4PFC	4PFC
	2	RHP	FATVAV	FATVAV
	3	VRV	VRV	VAV
	4	-	RHP	-
4	1	4PFC	FATVAV	VAV
	2	RHP	4PFC	FATVAV
	3	VRV	VRV	4PFC
	4	-	RHP	-
5	1	4PFC	4PFC	VAV
	2	RHP	FATVAV	4PFC
	3	VRV	RHP	FATVAV
	4	-	VRV	-
6	1	4PFC	FATVAV	VAV
	2	VRV	4PFC	FATVAV
	3	RHP	RHP	4PFC
	4	-	VRV	-

Table 8.4 Rank orders for systems established using MAVT

as 4PFC in case vignette 1, and so the analysis was concentrated on the more marginal decisions where there may be a change in rank. For first and second rank orders these were:

Engineer 1, Case Vignette 3

A change in preference from VAV to FATVAV requires a change in total weighted score of 90 points. Sensitivity analysis indicated that no such change in score was feasible.

Engineer 2, Case Vignette 2

A change in preference from 4PFC to FATVAV requires a change in total weighted score of 50 points. Sensitivity analysis indicated that this was just achievable given a simultaneous combination of changes in certain weights and scores, however the engineer judged that such coincidental changes were unfeasible, and was satisfied with the original decision for 4PFC.

Engineer 4, Case Vignette 2

A change in preference from FATVAV to 4PFC requires only a small change in total weighted score. A sensitivity analysis indicated that this was achievable in several ways. The model was therefore unable to indicate a definite preference.

Engineer 4, Case Vignette 3

A change in preference from VAV to FATVAV requires a change in total weighted score of 75 points. No such change was considered feasible during the sensitivity analysis.

Engineer 5, Case Vignette 3

A change in preference from VAV to 4PFC requires only a small change in total weighted score. The sensitivity analysis showed that this change could be achieved in a number of ways and so no definite preference could be indicated.

Engineer 6, Case Vignette 3

A change in preference from VAV to FATVAV requires a change in total weighted score of 75 points. The sensitivity analysis indicated that no change of this magnitude was feasible.

This sensitivity analysis was continued for the other rank orderings which appeared vulnerable to changes in weights or scores and the results are summarised in table 8.5 over. It was found that with the few exceptions indicated, the rank orderings were robust under sensitivity analysis.

Engineer	Case Vignette	Rank Orders Checked	Result
1	1 3	2/3 2/3	75 point change required - robust Sensitive to several changes
2	1 3	2/3 2/3	Sensitive to several changes 40 point change required - robust
3	1 2 3	2/3 3/4 2/3	30 point change required - robust 40 point change required - robust 40 point change required - robust
4	2 3	3/4 2/3	35 point change required - robust 45 point change required - robust
5	1 2 3	2/3 3/4 2/3	40 point change required - robust 100 point change required - robust Sensitive to several changes
6	1 2	2/3 3/4	80 point change required - robust 30 point change required - robust

Table 8.5 Result of sensitivity analysis for 2nd/3rd and 3rd/4th rank orders

8.7 ASSESSING THE VALIDITY OF THE COMPLETED MAVT MODEL

Up to this point, validity checks have been carried out at various stages in the construction of the MAVT models of each engineer in accordance with best practice outlined in section 7.3. This section now describes further tests on the validity of the completed models and discusses possible threats to the methodological validity of this stage of the study posed by facilitator effects.

Engineer	Rank	Case Vignette 1		Case Vignette 2		Case Vignette 3	
		Unaided	MAVT	Unaided	MAVT	Unaided	MAVT
1	1	4PFC	4PFC	4PFC	4PFC	FATVAV	VAV
	2	RHP	VRV	FATVAV	FATVAV	VAV	FATVAV
	3	VRV	RHP	RHP	RHP	4PFC	4PFC
	4	-	-	VRV	VRV	-	-
2	1	4PFC	4PFC	FATVAV	4PFC	4PFC	4PFC
	2	VRV	RHP	4PFC	FATVAV	FATVAV	FATVAV
	3	RHP	VRV	RHP	RHP	VAV	VAV
	4	-	-	VRV	VRV	-	-
3	1	4PFC	4PFC	4PFC	4PFC	4PFC	4PFC
	2	RHP	RHP	FATVAV	FATVAV	FATVAV	FATVAV
	3	VRV	VRV	RHP	VRV	VAV	VAV
	4	-	-	VRV	RHP	-	-
4	1	4PFC	4PFC	4PFC	FATVAV	FATVAV	VAV
	2	RHP	RHP	FATVAV	4PFC	VAV	FATVAV
	3	VRV	VRV	RHP	VRV	4PFC	4PFC
	4	-	-	VRV	RHP	-	-
5	1	4PFC	4PFC	4PFC	4PFC	VAV	VAV
	2	RHP	RHP	FATVAV	FATVAV	FATVAV	4PFC
	3	VRV	VRV	VRV	RHP	4PFC	FATVAV
	4	-	-	RHP	VRV	-	
6	1	4PFC	4PFC	FATVAV	FATVAV	FATVAV	VAV
	2	RHP	VRV	4PFC	4PFC	VAV	FATVAV
	3	VRV	RHP	RHP	RHP	4PFC	4PFC
	4	-	-	VRV	VRV	-	-

Table 8.6 Comparison of unaided preference judgements with MAVT judgements

8.7.1 Comparison with Previous Unaided Judgement

Section 8.1.2 described how each engineer produced a preference rank ordering for the air conditioning system options using unaided judgement. These unaided judgements were revealed and compared with those synthesised using MAVT. Table 8.6 on the previous page summarises this comparison by showing the unaided preference rank orderings for each engineer and each case vignette against the synthesised judgements of table 8.4.

The number of subjects is insufficient for a confident statistical analysis of the data (which would prove little in any case) however an inspection of the table reveals several interesting points. The change in rank where there is disagreement is never greater than one position (eg there is never a change in rank from first to third or fourth). When the table is considered in conjunction with the sensitivity analysis of section 8.6.2 it can be seen that without exception the rank changes occur in those cases where the decision is of a more marginal nature (highlighted by brackets in the table) although a more marginal decision is not always accompanied by a change in rank. The corollary of this of course is that agreements in rank correlate well with the choices which appear more obvious.

An important implication of this is that it does not appear from this data that any serious clashes between intuitive and synthesised judgement has occurred. To verify this, the engineers were asked to comment on any of the synthesised judgements which they felt were completely unacceptable (as opposed to debatable). In all cases they responded that the synthesised judgements were plausible and did not clash irreconcilably with intuition.

Given that the MAVT models seem valid to this extent, the more demanding and difficult question remains as to whether there is any indication that the synthesised judgements are superior to those produced by intuition. In an attempt to answer this question, the engineers were asked to consider those synthesised judgements which had produced a change in rank compared with the unaided judgements and to explain in detail why they were more confident in one or other judgement. For example, engineer 1 was asked about the changes in rank of VRV and RHP in case vignette 1, and of VAV and FATVAV in case vignette 3. He responded by saying that in case vignette 1, he had thought that RHP would have outperformed VRV to an extent that would have offset the capital cost difference. However an examination of the individual weights and scores for this case showed that RHP performed poorly on three important criteria when compared with VRV, namely *maintenance specialisation*, *load increase contingency* and *reduced load performance*. This had not been properly taken into consideration in his unaided judgement and so he believed that the MAVT ranking was correct. In case vignette 3, the extreme importance of *maintenance nuisance* is the main factor in swinging the analysis in favour of VAV. While the engineer had taken this criterion into consideration in his unaided judgement, he believed that he had underestimated its effect on the overall evaluation and agreed that the synthesised judgement was correct in this instance also. The above responses and those of the other engineers are summarised in table 8.7.

Engineer	Case Vignette	Rank Orders Discussed	Result
1	1	2/3	MAVT
	3	1/2	MAVT
2	1	2/3	Unaided
	2	1/2	MAVT
3	2	3/4	MAVT
4	2	1/2	MAVT
	2	3/4	MAVT
	3	1/2	MAVT
5	2	3/4	MAVT
	3	2/3	Unaided
6	1	2/3	MAVT
	3	1/2	MAVT

Table 8.7 Rankings upheld where disagreement occurred between unaided and MAVT synthesised judgements

As the table indicates, the engineers upheld the MAVT judgements on ten out of twelve occasions. On the two occasions where unaided judgement was thought superior, the sensitivity analysis of section 8.6.2 indicated that the synthesised judgements were highly sensitive and so the MAVT rankings were much weaker.

A further important implication of the foregoing is that the attributes which had to be omitted from the model due to their limited range of convenience (section 8.2.2), and the possible interaction between the attributes *load variations handled* and *reduced load performance* (section 8.4.1) did not adversely affect the synthesised judgements.

8.7.2 Discussion of Possible Facilitator Effects

As discussed in sections 7.1.2 and 7.4.2, the two major facilitator effects to which this study was thought vulnerable were forced consistency and a desire on the part of the subject engineers to please the analyst. In response to the first possible effect, an examination of the aided and unaided judgements summarised in table 8.6 indicates a healthy variation of opinion among the engineers generally. Moreover, there does not appear to be any greater agreement among the synthesised judgements than among the unaided judgements. Hence it does not appear as though an undesirable forced consistency effect was present in the overall judgements.

The second effect, the desire to please the analyst, is more difficult to detect and could have influenced the results of the study in two ways. Depending upon the subject engineer's interpretation of the aims of the study, he may have attempted (consciously or otherwise) to make his synthesised judgements agree or disagree with his previously made unaided judgements. Although such an effect is difficult to decern across only three cases, the variation in ranking agreements and disagreements for each engineer across the case vignettes suggests that no such effect was present to any significant extent. Furthermore, the complex and desegregated nature of the MAVT process is such that the deliberate 'fixing' of results would require the engineers to be possessed of an unfeasible mental agility and memory capacity.

The desire to please the analyst is more pernicious in connection with those assessments which rely on comments from the engineers, in particular, the question of whether the engineer believes the MAVT rankings to be superior to those obtained by unaided judgement. It is not possible to provide absolute protection against such an effect in this kind of study, however in each of the instances summarised in table 8.7 the engineers gave considered and plausible reasons for accepting one or other judgement. Criticism by the engineers of the MAVT technique was encouraged and at the conclusion of the study all had some reservations about the efficacy of the method (these are discussed in chapter 9). It was felt, therefore, that the responses of the subject engineers were honest and no serious bias was present.

8.8 CONCLUSIONS OF THE MAVT VALIDITY STUDY

In this study, the validity of MAVT was assessed in terms of the components of the model (attributes, value functions, weights etc) and in terms of the completed model (the final option rankings). In both respects, the results of this study are positive, but with some qualifications.

8.8.1 Validity of the MAVT Model Components

The first finding of the study, reported in section 8.2.2, was that certain attributes did not have a range of convenience which covered all of the options. As a result of this it had to be conceded even at this early stage that it would not be possible to develop an MAVT model which was *complete* in the decision analysis sense.

Notwithstanding this problem, when the MAVT process was continued, it was found that the value functions appeared valid, and that valid weights could be specified, although the large range in magnitude of the weights did suggest that some of the attributes were in fact constraints (section 8.5.2).

Additive difference independence was satisfied with a minor reservation on the part of two of the engineers, which did not appear to affect the validity of the model as assessed later.

Hence it was concluded that the model could not be a complete aid to the technical decision task, but within this limitation was valid insofar as there were no serious inconsistencies or failures in the components of the model.

8.8.2 Validity of the Completed MAVT Model

A sensitivity analysis of the completed models to reasonable uncertainties in weights and scores found that the MAVT rankings were robust with only a few exceptions. The MAVT rankings were then compared with those arrived at using unaided judgement, and it was found that there were several differences. An investigation of these differences established that disagreements in rank were associated with those MAVT rankings which were of a more marginal nature, while the rankings which appeared obvious from the MAVT analysis agreed well with the intuitive rankings. This evidence, and comments from the subject engineers themselves, indicated that there was no serious clash between the MAVT models and intuitive judgement.

The instances in which there were disagreements were examined in more detail, and in ten out of twelve cases the engineers believed that the MAVT ranking was superior to the intuitive ranking. In both the remaining instances, the sensitivity analysis had shown the MAVT rankings to be sensitive to changes in weights and scores and as such the rankings were weak. It is acknowledged that this assessment in particular is vulnerable to facilitator effects, however the examination of these in section 8.7.2 suggests that no serious bias was present.

The conclusion of this study therefore is that MAVT may be considered valid insofar as the MAVT models did prove more responsive than unaided judgement to information in the case vignettes.

Chapter 9

Summary of Conclusions and Discussion of Results

This chapter summarises the results and conclusions of the whole of this research, examines some practical issues concerning the application of MAVT which have emerged, and comments on areas which seem important for further research.

9.1 RESULTS AND CONCLUSIONS OF THE STUDY

Chapter 1 of this thesis set the context for this study of the validity of weighted scoring techniques as applied to the technical task of HVAC design. While attractive, it was explained that the weighted scoring technique as presented in the design methods literature and as recommended as an option appraisal technique in some practical design guides is an *a priory* method. It has reasonable face validity in that it seems to make sense, however there are few critical assessments of the assumptions inherent in the technique and no serious attempts to evaluate the validity of the technique as applied in a design context.

9.1.1 A Review of Design Theory

One of the fundamental assumptions underlying the application of weighted scoring techniques is the view that, at certain points, the design process can be characterised as a maximising decision making strategy wherein a number of design options are subjected to a careful evaluation in an attempt to identify the one which is best. This assumption arises out of the top-down, planned and methodical view of design as portrayed in the RIBA Plan of Work (RIBA, 1973) and the design methods literature in general. In particular, the RIBA Plan of Work and the congruent design guide for services engineers produced by BSRIA (1990) characterises the feasibility stage engineering services design task as one of assessing alternative engineering services design options.

However, a wider reading of the literature on design theory suggests that there is good reason to question the validity of even this fundamental assumption. In chapter 2, the theories of design as a process were reviewed. It was explained that the model of design assumed by the weighted scoring technique and embodied within the RIBA plan of work and similar guides is 'systematic'. Since the early 1970s, however, the systematic model of design has come under considerable criticism from an empirical point of view. Studies of how designs actually develop have repeatedly demonstrated that design does not follow the neat hierarchical approach of the systematic model, but rather is a more iterative and idiosyncratic process whereby an initial conjectured solution to one part of the design problem is progressively refined and checked for compatibility with other emerging aspects of the design. These criticisms were discussed and it was demonstrated that the tensions between writers in the field of design theory were largely explainable in terms whether the theories were intended to be prescriptive or descriptive. In conclusion, arguments were put forward in support of the view that a systematic approach to design, supported by appropriate methods, is still a realistic and worthwhile goal, provided the methods are adapted to the characteristics of the task to be performed.

9.1.2 The Characteristics of the Early Stage HVAC Design Task

According to the RIBA and BSRIA guides, the early stage HVAC design task can be characterised as the consideration of alternative HVAC solutions occurring at the feasibility stage. However, the review of chapter 2 cautions that it cannot be assumed that what is written in these publications represents the perception of practising engineers. Chapter 3 described how a preliminary study was carried out to determine whether there is a 'typical' form of design description for HVAC systems at the feasibility stage, and to provide insights into the process by which the designs evolve. This study found that while the management of technical design development appears on the surface to be very different to the top-down approach assumed by the systematic model, engineers do adopt a hierarchical process of HVAC design proceeding from the conceptual to the specific. The results further suggested that this process of design begins with a two stage appraisal in which the fundamental HVAC provision for each area of the building is assessed (eg natural ventilation, mechanical ventilation, air conditioning) and then an evaluation is made of generic system types from within this provision category. It was concluded that in the initial stages the 'typical' HVAC design process *can* be characterised as a decision between alternative solution types. This does not automatically imply that weighted scoring is an appropriate method for carrying out such an appraisal, however further investigation into the scope for the use of a formal decision aid such as the weighted scoring method seemed justified in this context.

Chapter 4 presented a review of the literature related to decision making and laid the philosophical basis for the continuing study. Models of the decision making process were discussed and the vexed question of how to assess the validity of a technique was considered.

9.1.3 The Validity of the Common Weighted Scoring Technique and a Further Exploration of the HVAC Design Task

In chapter 5, a detailed theoretical critique of the weighted scoring technique as it is normally described in the design methods literature was presented and it was shown that, even in a context-free sense, this simple version of the technique is theoretically deficient. This indicates that the indiscriminate application of such techniques should therefore be a cause for concern. It was also shown however, that a theoretically correct interpretation known as Multi Attribute Value Theory (MAVT) is possible.

MAVT, in common with other maximising decision making methods like Cost Benefit Analysis, makes certain important assumptions which mean that it is not an appropriate technique to apply in every decision situation. Chapter 6 related how a second study was carried out to gain a greater understanding of the characteristics of the decision making task in early stage HVAC design, with particular reference to how the structure of the decision situation conforms with the assumptions inherent in MAVT. The results of this study indicated that the task of deciding upon the fundamental HVAC provision for a space is not amenable to the application of methods like MAVT because the necessary assumptions are violated, but it was found that the more specific task of selecting a system type within a major provision category may be an appropriate problem for MAVT, with further research being most usefully concentrated on air conditioning system selection.

9.1.4 The Validity of MAVT in the Context of Air Conditioning System Selection

A third study was configured to assess the validity of MAVT in the context of air conditioning system selection for office buildings. Chapter 7 addressed the methodological difficulties associated with assessing the validity of a decision aid which were introduced in chapter 4, a problem to which there is no complete solution, and described the methods adopted here which it is hoped take some steps towards making an assessment of validity possible. The results of this study were reported in chapter 8.

It was concluded that it is not possible to develop an MAVT model which is a complete aid to the selection task because some aspects of the task violate the assumptions inherent in the technique. Models were finally constructed which had to compromise by omitting certain criteria, rendering some decision considerations external to the model. Nevertheless, the MAVT models were substantially if not totally complete, and so within this limitation it was found that MAVT can be considered a valid technique.

9.1.5 Further Discussion of the Results of Chapter 8

Part of the assessment of validity in this study was the extent to which the output of the MAVT models did not clash with intuitive judgements which the actors were confident in making, while at the same time exhibiting superior responsiveness to information about the decision task and thereby yielding improvements over unaided judgement. It was found that both of these qualities were present, however as several of the subject engineers pointed out, the improvements in judgement were evident only in marginal instances where no serious penalties would be incurred in selecting the next best alternative. This finding has implications for the usefulness of MAVT in practice. Although this research did not specifically address the issue of usefulness, having subjected the technique to rigorous scrutiny and applied it to a realistic problem it is felt that related to validity there is a dilemma in the use of the technique which must be confronted by the decision analysis community. It is believed that verifying validity requires the actor to have a reasonably well structured view of the decision situation in the first place. However, if the actor has such a view then it is unlikely that the construction of an MAVT model will produce any remarkable new insights. This was the situation in the air conditioning system selection problem investigated here. While the MAVT models developed are judged to be reasonably valid, it was argued by several of the engineers that their usefulness is debatable because they merely serve to confirm unaided judgements or improve upon them only in marginal instances.

The corollary of this is that decision analysis methods like MAVT would most beneficially be used in situations where the actor has a relatively poorly structured view of the decision situation. However this lack of structure implies that it would be exceptionally difficult if not impossible to verify the validity of the model in such cases, with the consequent risks involved to the integrity of the decisions made.

9.2 AREAS FOR FURTHER RESEARCH

The methods adopted for this research and the subsequent findings suggest certain areas which require further investigation.

9.2.1 Assessing the Validity of MAVT

More ways of assessing the validity of MAVT and decision aids in general must be investigated. It is evident that while some progress may be made towards an assessment of validity using the methods adopted in this research, more work is needed in developing definitions of validity which may then be tested empirically. For example, the results of this work indicate that models which are valid within understood limitations can be arrived at eventually by observing good elicitation and construction practices. Given that validity is ultimately achievable then, it is possible to define *more* valid techniques as those which converge on the final answer more quickly.

The severity of validity violations must also be taken into account. The 'validity dilemma' discussed above suggests that some validity requirements might be relaxed without serious errors being incurred. The extent of these errors might be possible to assess using sensitivity analysis in certain cases: for example errors in the specification of weights, or the effect of violations of additive difference independence could be assessed in this way.

9.2.2 The Elicitation of Attributes

The research reported here used the repertory grid technique to elicit attributes for use in the MAVT model. It was noted that this technique was found to be a highly efficient way of eliciting these attributes at the formative stages of the analysis, clarifying many points and averting later problems such as interactions which might lead to violations of additive difference independence. It is believed that more work is needed to investigate ways of eliciting a valid set of attributes, and to investigate in general the relationship between the essentially qualitative process of decision structuring and the quantitative MAVT technique. This would be particularly valuable in poorly structured decision situations. Work in this area is currently being undertaken in the Department of Management Science at the University of Strathclyde, where the use of cognitive mapping as an attribute structuring technique is being researched.

9.2.3 Incorporation of MAVT into the Wider Problem Solving Process

In chapter 4 it was emphasised that decision making *per se* is only one part of the wider and richer process of design problem solving. The study of a decision aid such as weighted scoring techniques was found to be justified in the context of HVAC design because according to the results of the preliminary study of chapter 3 and the literature on HVAC design, it seems that, typically, the early stage design of HVAC systems can be characterised as a choice between generic system types. It was suggested, however, that in more novel design situations the stages of analysis, synthesis and evaluation are highly interdependent. In such circumstances, MAVT will inform the other stages of the process in a way that was not possible to observe in this (or many other) studies in decision analysis. More research is therefore required on the application of MAVT and other decision analysis techniques in more novel and dynamic situations.

9.2.4 The Application of MAVT in Real Design Settings

It was made clear from the outset that this research would be concerned with the validity of weighted scoring techniques to the technical task of HVAC design in office buildings. This focus on the task was appropriate because had the technique proved invalid in this sense, there would have been little point in conducting more difficult assessments of validity amid the confusing and uncontrolled influences found in real settings. Now that the validity of MAVT has been established with some qualifications, it would be informative to apply the technique to the problem of HVAC design in a real setting in an effort to determine whether influences arising at the other interfaces depicted in figure 1.1 compromise the validity of the technique.

References

Ackoff, R L (1978) *The Art of Problem Solving.* Wiley, New York.

Adelman, L (1992) *Evaluating Decision Support and Expert Systems.* Wiley-Interscience, New York.

Adelman, L and Donnell, M L (1985) 'Evaluating decision support systems: A general framework and case study' in S Andriole (Ed) *Microcomputer Decision Support Systems* QED Information Sciences, Wellesley, MA.

Alexander, C (1966) 'A city is not a tree'. *Design.* Vol 20 pp46-55.

Anderson, N H (1982) *Methods of Information Integration Theory.* Academic Press, New York.

Anderson, N H and Zalinski, J (1988) 'Functional measurement approach to self estimation in multiattribute evaluation'. *Journal of Behavioural Decision Making.* Vol 1 pp191-221.

Andriole, S J (1989) *Handbook for the Design, Development, Evaluation and Application of Interactive Decision Support Systems.* Petrocelli, Princeton, NJ.

Archer, L B (1965) *Systematic Method for Designers.* The Design Council, London.

Archer, L B (1979) 'What became of design methodology?'. *Design Studies.* Vol 1 pp17-18.

Architektenkammer Nordrhein-Westfalen (Ed) (1986) *Architektenhandbuch 1986/87.* Dusseldorf.

ASCE (1988) *Manual of Professional Practice for Quality in the Constructed Project.* ASCE, New York

Asimow, M (1962) *Introduction to Design.* Prentice Hall, Englewood Cliffs, NJ.

Baron, J (1988) *Thinking and Deciding.* Cambridge University Press, Cambridge, MA.

Beach, L T and Mitchell, T R (1978) 'A contingency model for the selection of decision strategies'. *Academy of Management Review.* Vol 3 pp439-449.

Beesley, M E and Kain, J F (1964) 'Urban form, car ownership and public policy: an appraisal of "Traffic in Towns"'. *Urban Studies.* Vol 1 pp174-203.

Beheshti, R (1993) 'Design decisions and uncertainty'. *Design Studies.* Vol 14 pp85-95.

Belton, V (1985) 'The use of a simple multiple criteria model to assist in selection from a short list'. *Journal of the Operational Research Society.* Vol 36 pp265-274.

Belton, V (1986) *A Comparative Study of Methods for Multiple Criteria Decision Aiding.* PhD thesis, University of Cambridge.

Belton, V and Vickers, S (1990) 'Use of a simple multi-attribute value function incorporating visual interactive sensitivity analysis'. In C Bana e Costa (Ed) *Readings in Multiple Criteria Decision Aid.* Springer Verlag, Berlin.

Blanford, A E (1993) 'Applying the WOM to WOMBAT: evaluation of a tool to support learning about design evaluation'. *Design Studies.* Vol 14 pp228-246.

Borcherding, K, Eppel, T and von Winterfeldt, D (1991) 'Comparison of weighting judgments in multi-attribute utility measurement'. *Management Science.* Vol 37 pp1603-1619.

Bourgeois, L J and Eisenhardt, K M (1988) 'Strategic decision processes in high velocity environments'. *Management Science.* Vol 34 pp816-835.

Brown, R V, Kahr, A S and Peterson, C (1974) *Decision Analysis for the Manager.* Holt Rinehart and Winston, New York.

Brownlow, S A (1989) *Structuring Problems for Multi-Attribute Value Analysis.* PhD thesis, University of Cambridge.

Bruce, M (1993) Editorial. *Design Studies.* Vol 14 pp355-356.

BSRIA (1990) *Design Briefing Manual.* BSRIA Publications, Bracknell.

Buchanan, C (1963) *Traffic in Towns.* HMSO, London.

Buchanan, R (1990) 'Myth and maturity: towards a new order in the decade of design'. *Design Issues.* Vol 6 pp70-80.

Checkland, P (1981) *Systems Thinking, Systems Practice.* Wiley, Chichester.

Checkland, P and Scholes, J (1990) *Soft Systems Methodology in Action.* Wiley, Chichester.

Churchman, C W, Auerbach, L and Saden, S (1975) *Thinking for Decisions: Deductive Quantitative Methods.* Science Research Associates, Chicago, IL.

Corner, J L and Kirkwood, C W (1991) 'Decision analysis applications in the operations research literature, 1970-1989'. *Operations Research.* Vol 39 pp206-219.

Coshall, J T (1991) 'An appropriate method for eliciting construct subsystems from repertory grids'. *The Psychologist.* Vol 4 pp354-357.

Cross, N (1989) *Engineering Design Methods.* Wiley, Chichester

Darke J (1979) 'The primary generator and the design process'. *Design Studies.* Vol 1 pp36-44.

Delquie, P (1990) 'Inconsistent trade-offs between attributes: New insights in preference assessment biases'. *Transactions of the Operations Research Society of America/Institute of Management Sciences Joint National Meeting* Philadelphia, PA (unpublished).

DHSS (1987) *Option Appraisal: A guide for the National Health Service.* HMSO, London.

Dobias, A P (1990) 'Designing a mouse trap using the Analytic Hierarchy Process and Expert Choice'. *European Journal of Operational Research.* Vol 48 pp57-65.

Dockerty, J T (1985) 'Theory of Effectiveness Measures' *SHAPE Technical Centre Technical Memo STC TM-729* The Hague, The Netherlands.

Dyer, J S and Sarin, R A (1979) 'Measurable multi-attribute value functions'. *Operations Research.* Vol 22 pp810-822.

Eastman, C M (1970) 'On the analysis of intuitive design processes'. In G T Moore (Ed) *Emerging Methods in Environmental Design and Planning.* MIT Press, Cambridge, MA.

Edwards, W and Newman, J R (1982) *Multiattribute Evaluation.* Sage, Beverly Hills, CA.

Einhorn, H J (1971) 'Use of nonlinear, noncompensatory models as a function of task and the amount of information'. *Organisational Behaviour and Human Performance.* Vol 5 pp1-27.

Etzioni, A (1967) 'Mixed scanning: A third approach to decision making'. *Public Administration Review.* Vol 27 pp385-392.

Faludi, A (1986) *Critical Rationalism and Planning Methodology.* Pion, London.

Ferry, D J and Brandon P S (1984) *Cost Planning of Buildings.* (5th Ed) Collins, London.

Finch, E F (1989) *A Requisite Decision Model for the Selection of Mechanical and Electrical Services in Buildings.* PhD thesis, University of Reading.

Firlej, M and Hellens, D (1991) *Knowledge Elicitation: A practical handbook.* Prentice-Hall, Englewood Cliffs, NJ.

Fischer, G W (1979) 'Utility models with multiple objective decisions: do they accurately represent human performances?'. *Decision Sciences.* Vol 10 pp451-479.

Fishburn, P C (1965) 'Independence in utility theory with whole product sets'. *Operations Research.* Vol 13 pp28-43.

Fishburn, P C (1967) 'Methods of estimating additive utilities'. *Management Science.* Vol 13 pp435-453.

Fishburn, P C (1974) 'Lexicographic order, utilities and decision rules: a survey'. *Management Science.* Vol 20 pp1442-1471.

Flanagan, R, Norman, G, Meadows, J and Robinson, G (1989) *Life Cycle Costing: Theory and Practice.* BSP, London.

Fleming, M A (1973) 'Building Decisions and Economic Appraisal Techniques in Practice'. In G H Hutton and A D G Devonald (Eds) *Value in Building.* Applied Science Publishers, London.

Fransella, F and Bannister, D (1977) *A Manual for Repertory Grid Technique.* Academic Press, London.

Gass, S I and Torrence, S R (1991) 'On the development and validation of multicriteria ratings'. *Socio-Economic Planning Sciences.* Vol 25 pp133-142.

Gilleard, J (1988) 'The value of Value Engineering'. *Building Technology and Management.* Oct/Nov pp29-36.

Gilliland, S W, Schmitt, N and Wood I (1993) 'Cost-benefit determinants of decision process accuracy'. *Organisational Behaviour and Human Decision Processes.* Vol 56 pp308-330.

Glaser, B G (1978) *Theoretical Sensitivity: advances in the methodology of grounded theory.* Sociology Press, Mill Valley, CA.

Glaser, B G and Strauss, A L (1967) *The Discovery of Grounded Theory.* Aldine, Chicago, IL.

Grinold, R C and Marshall, K T (1977) *Manpower Planning Models.* Elsevier North-Holland, New York.

Gruber, T R (1989) *The Acquisition of Strategic Knowledge.* Academic Press, London.

Hall, A D (1962) *A Methodology for Systems Engineering.* Van Nostrand, Princeton, NJ.

Hart, A (1986) *Knowledge Acquisition for Expert Systems.* Kogan Page, London.

Hickling, A (1982) 'Beyond a linear iterative process?' In B Evans, J A Powell and R Talbot (Eds) *Changing Design.* Wiley, New York.

Hillier, W, Musgrove, J and O'Sullivan, P (1972) 'Knowledge and Design'. In W J Mitchell (Ed) *Environmental Design: Research and practice.* University of California, LA.

Hobbs, B F (1986) 'What can we learn from experiments in multiobjective decision analysis?'. *IEEE Transactions on Systems, Man and Cybernetics.* Vol 16 pp384-394.

Hobbs, B F, Chankong, V and Hamadeh, W (1992) 'Does choice of multicriteria method matter?'. *Water Resources Research.* Vol 28 pp1767-1779.

Hogarth, R M (1980) *Judgement and Choice.* Wiley, New York.

Hogarth, R M (1982) 'Discussion of an assessment of laboratory experiments in accounting'. *Journal of Accounting Research* Vol 20 pp108-116.

Hooker, C A, Leach, J J and McClennen, E F (Eds) (1978) *Foundations and Applications of Decision Theory, Vols 1 and 2.* D Reidel, Boston, MA.

Huber, G P (1986) *Managerial Decision Making.* Scott Foresman, Glenview, IL.

Humphreys P C, Wooler S and Phillips L D (1980) 'Structuring decisions: the role of structuring heuristics'. *Decision Analysis Unit Technical Report 80-1.* Brunel University (unpublished).

Howard, R A (1988) 'Decision analysis: practice and promise'. *Management Science.* Vol 34 pp679-695.

Jaccard J, Brinberg, D and Ackerman, L J (1986) 'Assessing attribute importance: a comparison of six methods'. *Journal of Consumer Research.* Vol 12 pp463-468.

Janis, I L and Mann, L (1977) *Decision Making.* Free Press, New York.

Johnson, L and Johnson, N E (1987) 'Knowledge elicitation involving teachback interviewing'. In A L Kidd (Ed) *Knowledge Acquisition.* Plenum Press, New York.

Jones, J C (1970) *Design Methods: Seeds of human futures.* Wiley, Chichester.

Jones, J C (1977) 'How my thoughts about design methods have changed during the years'. *Design Methods and Theories.* Vol 11 pp50-62.

Jones, J C and Thornley, D G (Eds) (1963) *Conference on Design Methods.* Pergamon Press, Oxford.

Keeney, R L (1981) 'Analysis of preference dependencies among objectives'. *Operations Research.* Vol 29 pp1105-1120.

Keeney, R L (1988) 'Building models of values'. *European Journal of Operational Research.* Vol 37 pp149-157.

Keeney, R L and Raiffa, H (1976) *Decisions with Multiple Objectives: Preferences and value trade-offs.* Wiley, New York.

Kelly, G A (1955) *The Psychology of Personal Constructs.* Norton, New York.

Kidd, A L (1987) *Knowledge Acquisition.* Plenum Press, New York.

Kirk, S J and Dell'Isola, A J (1983) *Life Cycle Cost Data.* McGraw-Hill, New York.

Klein, G A and Calderwood, R (1991) 'Decision models: some lessons from the field'. *IEEE Transactions on Systems Man and Cybernetics.* Vol 21 pp1018-1026.

Koestler, A J (1978) *The Act of Creation.* Random House, New York.

Kraus, R I and Mayer, J R (1979) 'Design: a case history'. In G T Moore (Ed) *Emerging Methods in Environmental Design and Planning.* MIT Press, Cambridge, MA.

Kuhn, T S (1970) *The Structure of Scientific Revolutions.* University of Chicago Press, Chicago, IL.

Langrish, J (1993) 'Case studies as a biological research process'. *Design Studies.* Vol 14 pp357-364.

Lawson, B (1980) *How Designers Think.* Architectural Press, London.

Layard, R (1972) *Cost-Benefit Analysis.* Penguin, London.

Legg, R C (1991) *Air Conditioning Systems Design, Commissioning and Maintenance.* B T Batsford Ltd, London.

Lera, S (1984) 'Architects design strategies: some justifications for current practice'. In J A Powell, I Cooper and S Lera (Eds) *Designing for Building Utilisation.* Spon, London.

Liebowitz, J and De Salvo, D A (Eds) (1989) *Structuring Expert Systems: Domain, design and development.* Yourdon, Englewood Cliffs, NJ.

Lock, A R (1983) 'Applying decision analysis in an organisational context'. In P Humphreys, O Svenson and A Vari (Eds) *Analysing and Aiding Decision Processes.* Proceedings of the 8th Conference on Subjective Probability, Utility and Decision Making. North Holland, Amsterdam.

Los, M (1981) 'Some reflections on epistemology, design and planning'. In M Dear and A J Scott (Eds) *Urbanisation and Urban Planning in Capitalist Society.* Methuen, London.

Luckman, J (1967) 'An approach to the management of design'. *Operational Research Quarterly.* Vol 18 pp345-358.

Mackinder, M and Marvin, H (1982) *Design Decision Making in Architectural Practice.* Research Paper 19, Institute for Advanced Architectural Studies, University of York.

Margolin, V (1989) *Design Discourse.* University of Chicago Press, Chicago, IL.

Markus, T A (1969) 'The role of building performance measurement and appraisal in design method'. In G Broadbent and A Ward (Eds) *Design Methods in Architecture.* Lund Humphreys, London.

Markus, T A (1973) 'Optimisation by evaluation in the appraisal of buildings'. In G H Hutton and A F G Devonald (Eds) *Value in Building.* Applied Science Publishers, London.

Martin, P Y and Turner, B A (1986) 'Grounded theory and organisational research'. *Journal of Applied Behavioral Science.* Vol 22 pp141-157.

Maver, T W (1970) 'Appraisal in the building design process'. In G T Moore (Ed) *Emerging Methods in Environmental Design and Planning.* MIT Press, Cambridge, MA.

Maver, T W (1979) 'The concept of modelling in architectural design'. *Design Methods and Theories.* Vol 13 pp67-82.

Miller, D W and Starr, M K (1967) *The Structure of Human Decisions.* Prentice-Hall, Englewood Cliffs, NJ.

Miller, G A (1956) 'The magical number seven, plus or minus two'. *Psychological Review.* Vol 63 pp81-97.

Mittal, S and Dym, C L (1985) 'Knowledge acquisition from multiple experts'. *AI Magazine*. Summer, pp32-37.

Montgomery, H (1983) 'Decision rules and the search for a dominance structure: Towards a process model of decision making'. In P Humphreys, O Svenson and A Vari (Eds) *Analysing and Aiding Decision Processes.* North Holland, Amsterdam.

Montgomery, H and Svenson, O (1976) 'On decision rules and information processing strategies for choices among multi-attributed alternatives'. *Scandanavian Journal of Psychology*. Vol 17 pp283-291.

Moselhi, O and Deb, B (1993) 'Project selection considering risk'. *Construction Management and Economics*. Vol 11 pp45-52.

Nadler, G (1981) *The Planning and Design Approach.* Wiley, New York.

Neuman, V A and Guven, H (1988) 'A simple ranking method for choosing HVAC systems'. *Energy Engineering*. Vol 85 pp28-35.

Newall, A and Simon, H A (1972) *Human Problem Solving.* Prentice-Hall, Englewood Cliffs, NJ.

Newland, P, Powell, J A and Creed, C (1987) 'Understanding architectural designers' selective information handling'. *Design Studies*. Vol 8 pp2-16.

Newman, J R (1975) *Assessing the reliability and validity of multiple attribute utility assessment procedures*. Social Science Research Institute, University of Southern California, Los Angeles, CA.

Nisbett, R E and Wilson, T D (1977) 'Telling more than we can know: verbal reports on mental processes'. *Psychological Review.* Vol 84 pp231-259.

Oberstone, J (1990) *Management Science: concepts, insights and applications.* West Publishing Company, New York.

Obgorn, J M and Johnson, L (1984) 'Conversation theory'. *Kybernetes* Vol 13, pp177-181.

Olshavsky, R W (1979) 'Task complexity and contingent processing in decision making: a replication and extension'. *Organisational Behaviour and Human Performance*. Vol 24 pp300-316.

Pask, G (1974) *Conversation, Cognition and Learning: A cybernetic theory and methodology*. Elsevier, London.

Patton, Q M (1980) *Qualitative Evaluation Methods.* Sage Publications, Beverly Hills, CA.

Payne, J W (1982) 'Contingent decision behaviour'. *Psychological Bulletin.* Vol 92 pp 382-402.

Payne, J W, Bettman, R J and Johnson, E J (1988) 'Adaptive strategy selection in decision making' *Journal of Experimental Psychology: Learning, Memory and Cognition.* Vol 14 pp534-552.

Pearce, D W and Nash, C A (1981) *The Social Appraisal of Projects: a text in cost-benefit analysis.* MacMillan, London.

Pearl, J (1984) *Heuristics.* Addison-Wesley, New York.

Phal, G and Beitz, W (1984) *Engineering Design.* The Design Council, London.

Phillips, L D (1984) 'A theory of requisite decision models'. *Acta Psychologica.* Vol 56 pp29-48.

Piper, R J (Ed) (1987) *The Architect's Handbook of Professional Practice.* AIA Document B 162, AIA Press.

Pitz, G F, Heerboth, J and Sachs, N J (1980) 'Assessing the utility of multi-attribute utility assessments'. *Organisational Behaviour and Human Performance.* Vol 26 pp65-80.

Podaskoff, P M and Dalton, D R (1987) 'Research Methodology in Organisational Studies'. *Journal of Management.* Vol 13 pp419-441.

Popper, K R (1963) *Conjectures and Refutations.* Routledge and Kegan Paul, Oxford.

Popper, K R (1972) *Objective Knowledge.* Oxford University Press, Oxford.

Poulton, E C (1979) 'Models of biases in judging sensory magnitude'. *Psychological Bulletin.* Vol 86 pp777-803.

Powell, J A and Russell B (1983) 'Model blindness in design'. In A Pendrith and G de Zeeuw (Eds) *Problems of Levels and Boundaries.* Princelet Editions, London.

Powell, J A, Thompson, M and Wynne B (1984) 'Cerebral hygiene and information drips'. In J A Powell, I Cooper and S Lera (Eds) *Designing for Building Utilisation.* Spon, London.

Quade, E (1975) *Analysis for Public Decisions.* Elsevier, New York.

Quine, W V (1976) *The Ways of Paradox and Other Essays.* Harvard University Press, Cambridge, MA.

Raiffa, H (1968) *Decision Analysis.* Addison-Wesley, Reading, MA.

Ranyard, F H (1977) 'Risky decisions which violate transitivity and double cancellation'. *Acta Psychologica.* Vol 41 pp449-459.

RIBA (1973) *Plan of Work for Design Team Operations.* RIBA Publications, London.

Riedel, S L and Pitz, G F (1986) 'Utilisation orientated evaluation of decision support systems'. *IEEE Transactions on Systems, Man and Cybernetics.* Vol 16 pp980-996.

Ritell, H W and Webber, M M (1973) 'Planning problems are wicked problems'. *Policy Sciences.* Vol 4 pp155-169.

Roozenburg, N F M and Cross, N (1991) 'Models of the design process: integrating across the disciplines'. *Design Studies.* Vol 12 pp215-220.

Sage, A P (1981) 'Behavioral and organisational considerations in the design of information systems and processes for planning and decision support'. *IEEE Transactions on Systems, Man and Cybernetics.* Vol 11 pp640-678.

Sage, A P (1986) 'An overview of contemporary issues in the design and development of microcomputer decision support systems'. In S J Andriole (Ed) *Microcomputer Decision Support Systems.* QED Information Sciences, Wellesley, MA.

Sainfort, F C, Gustafson, D H, Bosworth, K and Hawkins, R P (1990) 'Decision support systems effectiveness: conceptual framework and empirical evaluation'. *Organisational Behaviour and Human Decision Processes.* Vol 45 pp232-252.

Sanoff, H (1977) *Methods in Architectural Programming.* Dowden Hutchinson and Ross, Pitsburg, PA.

Sassone, P G and Schaffer, W A (1978) *Cost-Benefit Analysis: a handbook.* Academic Press, New York.

Schoemaker, P J H and Waid, C D (1982) 'An experimental comparison of different approaches to determining weights in additive utility models'. *Management Science.* Vol 28 pp182-196.

Scott, A J and Roweis, S T (1977) 'Urban planning theory and practice: A reappraisal'. *Environment and Planning.* Vol 9 pp1097-1119.

Self, P (1970) 'Nonsense on stilts. Cost-benefit analysis and the Roskill Commission'. *Political Quarterly.* Vol 41 pp249-260.

Self, P (1975) *Econocrats and the Policy Process: the politics of cost-benefit analysis.* MacMillan, London.

Shaw M L and Gaines B R (1987) 'An interactive knowledge elicitation technique using personal construct theory'. In A L Kidd (Ed) *Knowledge Acquisition*. Plenum Press, New York.

Silverman, B G (1991) 'Expert critics: Operationalising the judgement/decision making literature as a theory of bugs and repair strategies'. *Knowledge Acquisition*. Vol 3 pp175-214.

Simon, H A (1957) *Administrative Behaviour*. Macmillan, New York.

Simon, H A (1973) 'The structure of ill-structured problems'. *Artificial Intelligence*. Vol 4 pp181-200.

Simon, H A (1975) 'Style in design'. In C M Eastman (Ed) *Spatial Synthesis in Computer-aided Building Design*. Applied Science Publishers, London.

Smith, H W (1981) *Strategies of Social Research*. Prentice-Hall, Englewood Cliffs, NJ.

Stanford, H W (1988) *Analysis and Design of Heating, Ventilating and Air Conditioning Systems*. Prentice Hall, Englewood Cliffs, NJ.

Stern, G J A (1976) 'SOSIPing, or Sophisticated Obfuscation of Self Interest and Prejudice'. *Operational Research Quarterly*. Vol 27 pp915-929.

Stewart, T R and Ely, D W (1984) 'Range sensitivity: a necessary condition and a test for the validity of weights'. *National Centre for Atmospheric Research Technical Report 3141-84/14*. Boulder, CO.

Stewart, V and Stewart, A (1981) *Business Applications of Repertory Grid*. McGraw-Hill, London.

Stillwell, W G F, von Winterfeldt, D and John, R S (1987) 'Comparing hierarchical and non-hierarchical weighting methods for eliciting multi-attribute value models'. *Management Science*. Vol 33 pp442-450.

Stewart, V and Stewart, A (1981) *Business Applications of Repertory Grid*. McGraw-Hill, London.

Sugden, R and Williams, A (1978) *The Principles of Practical Cost-Benefit Analysis*. Oxford University Press, Oxford.

Svengren, L (1993) 'Case study methods in design management research'. *Design Studies*. Vol 14 pp444-456.

Svenson, O (1977) 'Intuitive extrapolations of a growth function'. *Scandinavian Journal of Psychology* Vol 18 pp339-344.

Takeda, H, Hamada, S, Tomiyama, T and Yoshikawa, H (1990) 'A cognitive approach to the analysis of design processes'. In J R Rinderle (Ed) *Design Theory and Methodology - DTM'90.* American Society of Mechanical Engineers, Chicago, pp153-160.

Townsend, R (1970) *Up The Organisation.* Knopf, New York.

Turner, B A (1981) 'Some practical aspects of qualitative analysis: one way of organising the cognitive processes associated with the generation of grounded theory'. *Quality and Quantity.* Vol 15 pp225-247.

Tversky, A (1972) 'Elimination by aspects: a theory of choice'. *Psychological Review.* Vol 79 pp281-299.

Ullman, D G, Deitterich, T G and Stauffer, L A (1988) 'A model of the mechanical design process based on empirical data'. *AI EDAM.* Vol 2 pp33-52.

Van Gundy, A B (1988) *Techniques of Structured Problem Solving.* Van Nostrand Reinhold, New York.

VDI-Richtlinie 2222 Blatt 1 (1973) *Konzipieren technischer Produkte.* VDI-Verlag, Dusseldorf.

von Winterfeldt, D and Edwards, W (1986) *Decision Analysis and Behavioral Research.* Cambridge University Press, Cambridge.

Wagner, H M (1975) *Principles of Management Science.* Prentice-Hall, Englewood Cliffs, NJ.

Warszawski, A (1984) 'Evaluation of building design alternatives'. In *CIB W55 3rd International Symposium on Building Economics, Vol 2.* National Research Council of Canada, Ottawa.

Waterman, D A (1986) *A Guide to Expert Systems.* Addison-Wesley, Reading, MA.

Watson, P C and Evans, J (1975) 'Dual processes in reasoning?' *Cognition.* Vol 3 pp141-154.

Watson, S R (1981) 'Decision analysis as a replacement for cost-benefit analysis'. *European Journal of Operational Research.* Vol 7 pp242-248.

Watson, S R and Buede, D M (1987) *Decision Synthesis.* Cambridge University Press, Cambridge.

Wierzbicki, A P (1983) 'Critical essay on the methodology of multiobjective analysis'. *Regional Science and Urban Economics.* Vol 13 pp5-29.

Willem, R A (1991) 'Varieties of design'. *Design Studies*. Vol 12 pp132-136.

Wooler, S (1984) 'Let the decision maker decide! A case against assuming common occupational value structures'. *Decision Analysis Unit Technical Report 84-4* London School of Economics, London.

Yin, R (1989) *Case Study Research: design and methods.* Sage Publications, CA.

Appendix 1

Descriptions of Case Studies in the Preliminary Study of Chapter 3

Project 1

A corporate headquarters building of 3,500 sq m in a green field site. The building was to house the company's central administration facilities and would comprise general office areas, private offices for senior staff, meeting rooms, a cafeteria and a special audio-visual room for presentations and teleconferences.

There were general plans of 1:200 scale, clearly indicating the plan shape of the building. The arrangement of the open plan office spaces had been determined, and there were notes to the effect that these open plan areas were to be flexible, with the option of introducing modular partitioning if required. An area on one floor had been allocated for the private offices and meeting rooms, but without specifically indicating the partitioning between them. The audio-visual room was however clearly marked on the drawings.

Circulation/service cores housing lifts, stairs, toilets and plant rooms were indicated by rectangular blocks on the plans and, with the exception of the main entrance and reception area, there were no details of how these areas would be laid out. The cafeteria was similarly indicated as a block on the plans.

An approximate structural grid had been roughed out for the office areas. The location of beams and columns were indicated by lines on the plans and annotated with the expected size of structural members. There was no such information for the service areas or the cafeteria.

There were 1:20 scale detailed sections through certain parts of the open plan office areas showing the proposed HVAC, electrical and lighting system arrangements. Considerable written information describing the reasons behind the selection of the HVAC system types had been produced in preparation for making a presentation to the client. Leading up to this there had been some discussion about the amount of glazing on the southerly facade, and the use of various solar shading devices had been considered. Computer thermal simulations had been carried out to analyse the performance of the open-plan office areas with respect to maximum temperatures likely to be achieved in spring and summer with and without air conditioning. The results of these simulations had served to confirm that the engineer's original choice of HVAC system for the open plan spaces was appropriate, although much leeway had to be allowed as the simulation took place ahead of detailed design by the architect of the building external walls and roof.

There was also a 1:20 detailed plan of the audio-visual room showing the proposed furniture layout together with a wealth of information on the audio-visual services including wiring schematics and specimen equipment specifications.

There was no detail on the central or secondary plant rooms or the services for areas out with the office spaces. At this stage, the plant room areas had been estimated using a unit area allocation based on similar previous jobs, and the type of HVAC provision for the canteen area had been noted but without any accompanying design information.

Project 2

A corporate headquarters building of 8,000 sq m in a green field site. The building was required to re-locate the company's main administration departments and a special computer centre. The areas to be accommodated were general open-plan office space, private executive offices, meeting rooms, a seminar room, the computer centre, a staff canteen and some ancillary spaces such as a mail room.

There were 1:200 free hand drawings of the building plan, with a rough structural grid marked out for the main blocks but not for the single storey extensions which were to house the canteen and some of the central plant. There were also sketched elevations on the same scale. Areas had been marked off for the office spaces, the computer centre, the staff canteen and some of the ancillary areas. At this stage, no differentiation in type of office space (ie open plan or private) had been made. The location of open stairways and lifts had been marked on the plans, and a rough area had been set aside on each floor for secondary plant space and toilets.

A 'typical' 1:20 scale section through a main block had been sketched freehand to show how the HVAC system for the open plan office areas would be accommodated, showing ducts, air supply and extract terminals and coordination with the proposed lighting scheme. There had been some debate over the choice of HVAC system for the open plan office areas. The client in this case had certain vested interests which caused them to favour one particular system type. However, the HVAC engineer did not consider that system appropriate for such a building and had to make a very persuasive case for proposing a different system. This involved producing a three page option appraisal of six different HVAC systems, and demonstrating that, on a number of criteria, the system proposed by the engineer was the most suitable.

There were similarly detailed 1:20 sketches of feature atria which were to link the main blocks of the building together. There had been detailed consideration of the solar heat gains due to these atria links, and computer thermal simulations of this particular architectural feature had been carried out. As a result, consideration was being given to the use of automatic blinds to reduce peak summertime gains.

There was no information on main or secondary plant rooms other than the space allocated for them, nor was there any detail of the design of the computer centre, canteen area or other ancillary spaces. Checks had been made, however, that plant space was adequate and that there was sufficient space for services in the computer centre, canteen etc by assuming a particular type of HVAC system and using broad rules of thumb provided by an in-house design guide and data book produced by the consultancy's own research and development department.

Project 3

A speculative office development of 50,000 sq m in a city centre site. The building was to be a multi-storey deep plan construction with a plan shape that was constrained by the site. In addition to the open plan office areas, there was to be a restaurant, basement car parking and a sub-basement below this for central plant rooms.

There were 1:100 scale plans for a typical floor with a structural grid shown, and 1:200 sketched elevations. Stairways, lifts and toilet facilities were to be accommodated in towers located at various points around the perimeter, and the position and area of these had already been fixed in accordance with the developer's rules.

The roof scape of the building had also been the subject of some detailed consideration by the architect. Planning restrictions prevented the location of exposed plant on the roof, and so some discussion had taken place regarding the siting and screening of roof located plant.

A certain amount of secondary plant room space was required for each floor, and this would need to be at the perimeter for access to an outside wall (for fresh air intakes and exhausts). Although the broad total areas per floor were known, at this point the number and locations of the plant rooms were still the subject of negotiation with the developer. An approximate area had been allocated for the dining facilities and the kitchens, and the general level of HVAC servicing envisaged for these areas had been noted, but there was no detail beyond this.

Larger scale sections through a floor of the building had not yet been drawn, however contingent allocations for the height of the raised floor and the depth of suspended ceiling void had been agreed based on a 'normal' provision for such a building. This was possible even although little design work had been carried out because the range of HVAC systems which could have been used had already been constrained as a result of the developer's brief which specifically called for the use of one of only two popular system types. Rules of thumb coupled with the engineer's experience had indicated that the space allocations would be sufficient for either option. Nevertheless, the project engineer was concerned that certain details such as the distribution of major duct work from secondary plant rooms could cause problems in relation to the ceiling space allowance depending upon where these plant rooms were finally located. More detailed sections would be drawn up for both of the possible HVAC system types as soon as the issue of plant room locations had been resolved, and computer simulations to test the predicted performance of both HVAC system types were planned for the near future to assist in the final selection.

Two issues had already been the subject of detailed consideration by the engineer, however. Firstly, the developer's brief had originally called for the HVAC system to be capable of offsetting gains from lighting and office equipment of 150 W/sq m, however the consulting engineer had contested this as being excessive. Secondly, the developer was interested in the use of an ice storage system to take advantage of off-peak electricity for the air conditioning as they believed this would present a 'greener' image. (The consulting engineer pointed out that while there was some 'green' value in the use of off-peak electricity inasmuch as it reduced peak time loads and so the required generation and transmission capacity, it did not actually reduce the amount of fuel used.)

The consulting engineer spent several weeks investigating comparable existing buildings, both to learn about ice storage and to assess the real gains from lighting and equipment. As a result it was concluded that ice storage could be incorporated into the

HVAC design and that the gains from lighting and equipment should be assumed to be 55 W/sq m with additional capacity for local cooling of 15 W/sq m in a limited number of areas.

Project 4

An 8,000 sq m office building on the outskirts of a city. The building was being constructed by a developer for a government department so that the departmental staff could be re-located from several older existing buildings within the city. The building was to accommodate general office space, private offices, meeting rooms and a staff canteen.

1:100 scale plans of the ground floor and a typical upper floor of the building (which contained a central atrium) had been sketched. These plans illustrated a rough structural grid with approximate beam and column dimensions. The areas for the office space had been delineated, although at this stage there was no distinction between general open plan space and private offices. A rough area allocation for circulation/service cores and the canteen had been made and the expected locations for these had been indicated on the plans. These areas had been determined from norms, however, and no consideration had yet been given to layout.

The government department who would occupy the building was particularly concerned about the environmental image of the building, and had requested that it not be air conditioned. The consulting engineers had accordingly spent much time investigating various forms of mechanical ventilation which would make use of the structure of the building to provide a damping effect against peak summertime temperatures. They had also assessed various methods of solar shading for the windows. 1:20 scale detailed sections through an open plan office area had been produced showing alternative mechanical ventilation arrangements with their associated ceiling and/or floor void depth requirements and also sketched details of solar shading devices. The most promising of these proposed designs was selected for computer simulation which was carried out on behalf of the consulting engineers by a specialist company. The results indicated that this design in conjunction with external solar shades could hold the temperature in the open plan office areas down to an acceptable level, albeit considerably higher than would normally be tolerated in a fully air conditioned building. It was decided, however, that the private offices and meeting rooms should be provided with a small and flexible air conditioning system. The canteen would also be air conditioned by a self contained packaged unit. Beyond the selection of the generic system type however, no details of these systems had been considered at this point.

Project 5

A 6,000 sq m integrated circuit manufacturing works in a new town industrial park. The building was to comprise a factory with an attached block of 1,500 sq m which would accommodate offices and a canteen.

There were 1:200 scale plans for the building with a sketched structural grid for the office block but only vague structural information for the factory. The office areas were marked on the drawings, as was an area for the canteen. It was noted that the services for the office areas were to be flexible so that modular partitioning could be used where required. An allowance had been made in the total floor area of this block for circulation and toilet facilities, but these had not yet been located on the drawings. In the factory, areas had been marked off for certain production processes and services plant rooms had been drawn in.

In this case, the client had a briefing and design guide of their own which contained some highly specific requirements for the design of both production and office spaces. Most of the work space allocations came directly from this guide. The guide also stipulated the type of HVAC system to be used in the various spaces, and while the consulting engineers accepted most of these requirements for the production spaces (as the client had more experience in these highly specialised areas) they had successfully challenged some of the guide's requirements for the office areas.

This challenge had involved going into some detail about the merits and de-merits of various HVAC systems, and persuading the client that, in this case, a different system would be more suitable.

There were sketched 1:50 scale sections of an open plan office area illustrating how the HVAC services would be accommodated. There were also similar sections through certain areas of the factory where the main HVAC duct routes would pass. There was no information regarding the canteen other than the area allocation made on the plans and a note that the suspended ceiling void would be increased to a depth of one metre in this area to house the duct work for a seperate packaged constant volume air conditioning system.

Project 6

A private printing facility of 1,000 sq m for a large financial institution. The building was to be constructed as an extension to an existing building and would be designed to provide storage for paper and inks, accommodate the printing presses and cutting and binding machinery, and provide temporary storage for the finished product.

General plans of 1:100 scale had been produced with tentative information about the portal frame design of the building (which was to be a basic industrial style shed). The plan indicated the layout of the storage and production areas, which had actually been determined by the client to facilitate efficient production. The plan also indicated the location of the services plant room.

In addition to the storage and production areas, the client had specified the requirement for certain ancillary areas such as washrooms, a small kitchenette/dining area and an office, however the location of these spaces was left to the discretion of the design team. At this point, although a gross area allowance had been made for these spaces, their arrangement had not been discussed.

In the brief, the client had requested that the building rely on a mechanical ventilation system only, rather than be air conditioned. The consulting engineer carried out a simple calculation which demonstrated that mechanical ventilation alone could not cope with the heat gains from the printing machinery, nor could it maintain the relative humidity within the required limits for satisfactory storage and production. The client conceded that air conditioning was indeed required, however they made it clear that they were concerned about the attendant cost implications.

With this sensitivity to cost in mind, the consulting engineer proceeded to investigate the appropriate forms of HVAC system. The system requirements were left entirely to the engineer in this case. Although the client organisation did have other printing facilities, they had no firm guidelines on the servicing and environmental requirements of the machines. The engineer therefore had to investigate temperature, humidity and ventilation recommendations for the various processes prior to proceeding with the system selection.

For the production areas, the engineer narrowed the range of system options down to two, and produced some detailed design information on both options. 1:20 sections through certain parts of the production area had been drawn, indicating the HVAC ducting arrangements under both options. Main duct routes leading from the plant room had also been sketched on the plans, the duct sizes being "educated guesses".

There was no detail on the services for areas out with the production space except notes to the effect that the storage areas would be served by the same system as the production space and the ancillary areas would be naturally ventilated (being next to an outside wall) and would be provided with a form of electric heating system.

Project 7

A laboratory building of 1,800 sq m in an industrial/research park. The building was to accommodate dissection laboratories, a number of animal keeping rooms, specialist areas such as embryo incubation rooms, tissue culture rooms and sterilisation facilities, staff offices, a small reference library, a staff kitchenette and hygienic staff changing rooms.

There was an extremely detailed brief from the client regarding the requirements for some of the areas in the building. In particular, there were strict stipulations on the adjacencies of certain areas to ensure that no contaminated materials could affect the results of the highly sensitive experiments being conducted. Entry to the building had to be strictly controlled, and the staff had to observe a hygiene procedure in moving from one area to another.

These considerations led to the production of area layout drawings at an early stage in the design. The architect had supplied 1:100 scale plans of the building with an approximate structural grid. The area layouts had been determined for some areas but not for others. On the ground floor, the building entrances and exits had been located, together with the staff changing rooms, sterilisation facilities, quarantine areas and animal keeping rooms. On the upper floor, the location of the tissue culture rooms and embryo incubation rooms had been fixed. There were also areas tentatively set aside for toilets, a stair well and a lift. The remaining areas had not been fixed, however positions had been provisionally marked off for the services plant rooms on the ground floor. This was principally because the plant rooms had to face into the secure courtyard of the building so that there were no air intake louvres on the outside facade. The clients feared action by animal rights activists, and this security consideration had influenced other design decisions affecting the outer facade such as there was to be no externally run duct work and no windows at ground level.

In addition to the above area, adjacency and security requirements, the brief also contained details on the HVAC provision for some of the areas. The HVAC specifications for the animal keeping rooms and the tissue culture laboratories were particularly onerous. These areas required tightly controlled environmental conditions, very high ventilation rates and low noise levels. In addition, the tissue culture laboratories had fume cupboards which would be used to handle toxic gasses. To satisfy planning regulations on the design of the extract arrangements for these cupboards, the engineer had been required to commission wind tunnel tests using a scale model of the site.

The consulting engineer had produced 1:20 scale sections showing the proposed design of an animal keeping room and a tissue culture room, with sizes of air ducts and pipes. There were also dimensioned sketches indicating the arrangement of the fume cupboard exhaust flues on the roof, and the route for ductwork from roof mounted air handling units down to the ground floor animal keeping rooms.

There was no detail on the services for the other areas, but it was noted that they had more routine HVAC requirements. The offices were to be naturally ventilated with a panel radiator heating system, and most of the other areas without opening windows would require only normal mechanical ventilation.

Project 8

A library annexe of 1,500 sq m in a city centre gap site, accommodating open plan areas for books, journals etc interspersed with reading tables, a service desk and adjacent staff office area, photocopying facilities for readers, an archive storage area, a microfiche/microfilm reading room, a bank of audio-visual cubicles for the viewing of video tapes and a coffee room.

There were 1:100 scale plans with a reinforced concrete structural grid roughed out. The shape of the building had effectively been determined by the nature of the gap site. The position of the main entrance and the service desk/staff office area had been decided, as had the position of the main stair well and lift shaft. In the basement, an area had been set aside for a plant room with the remainder to be used for archive storage. The positioning of the service desk, staff office area and stair well had all been dictated by the location of the main entrance for security reasons, and the client's brief had been most specific about these adjacency requirements. The upper floors had no specific layout information save the location of vertical services ducts running from the roof which had been sketched in, and the positioning of the other areas such as the microfiche/microfilm reading room had yet to be decided.

A 1:100 plan of the roof showing in some detail the arrangement of the air handling units and air cooled condensers for the chiller housed in the basement plant room had been sketched by the consulting engineer. Notes in the file indicated that there had been detailed consideration over the unfortunately tight positioning of the air handing units to ensure that no 'short circuiting' occurred (ie where the exhaust air grille of one unit is so close to the fresh air intake of another that stale exhaust air is immediately sucked back into the building again). The notes also showed that thought had been given to the location of the chiller on the roof, but that this had been discounted on the grounds of noise.

1:20 sections through the open plan floor area had been sketched, showing the suspended ceiling void space for services and how the main HVAC ducts would branch off to distribute air to the area. These sections were extremely detailed, with duct sizes, bend radii and other dimensions marked on the drawings.

Appendix 2

An Example Concept Note and Theoretical Memo as used in Grounded Theory

CONCEPT NOTE 11

Choice of areas for detailed HVAC design

(extract)

(3) IN 2/1[1] Re 1:20 plan of AV room with possible furniture layout and a detailed description of the AV services (see CN 10)[2]. "The MD was over in Japan seeing one to the customers they [the client] export to, and they [the Japanese] had this AV room where they did presentations and held teleconferences. He was really impressed with this and came back saying that their new HQ had to have one of these. They made a big thing of it in the initial briefing, so we thought we'd better take that on board." (see also CN 15).

(4) IN 4/2 Re 1:20 section, open plan office areas (see CN 10). "You've got to know what you're doing in these spaces. We'd eventually persuaded them [the client] that in-ceiling VAV was the best system for the kind of deep plan open plan offices they wanted...but the architect had this lighting scheme idea in mind for the upper floors [uplighters illuminating a pine clad barrel vault ceiling] and so I had to be sure that we could run the ducts and conceal them without spoiling the effect.

(5) IN 4/4 The client had specifically requested that the building not be air conditioned (CN 15). Given the gains to the open plan space, however, admittance and response factor calculations on HEVACOMP indicated that peak temperatures in these areas would reach 29-30 deg C even with high volume ventilation. This suggested that a number of passive measures would be required in addition to the ventilation. The engineer sketched four possible ways (free hand 1:20) in which as much of the building structure as possible could be exposed (see CN 10).

(6) IN 3/7 Re 1:20 sections of animal keeping room (CN 10). "These sections were virtually the first thing we drew on this job. They needed 23 deg C + or - 1 and 15 air changes with near laminar flow. They also wanted NR 30 because apparently mice are noise sensitive animals. It wasn't going to be easy...and there were twelve of them [rooms] so we had to get it right...We also had to get the supply and extract ducts into the corridor somehow and I wasn't sure if we could do it inside the architect's floor to ceiling height."

---

1 Number in brackets denotes note number. Code following refers to source in original interview notes.

2 Cross references to other concept notes.

THEORETICAL MEMO 5

Detailed design as an investigative process at the feasibility stage

---

CN 11 notes 4-7, 10, 12-15 suggest that the more detailed design was carried out to *investigate* certain points.

Note 4 indicates that if the concealment of the ducts had not been an important detail, the engineer would not have been concerned about certain aspects of the distribution duct work design.

Note 5 shows how the incorporation of passive measures into the building and HVAC design could not be assumed - some more detailed design was needed to see if and how this could be accomplished.

In note 6 the onerous HVAC requirements for the animal rooms created a similar problem. The engineer did not know if and how it could be done, and so some design work was necessary to work out important details that could affect the architectural design.

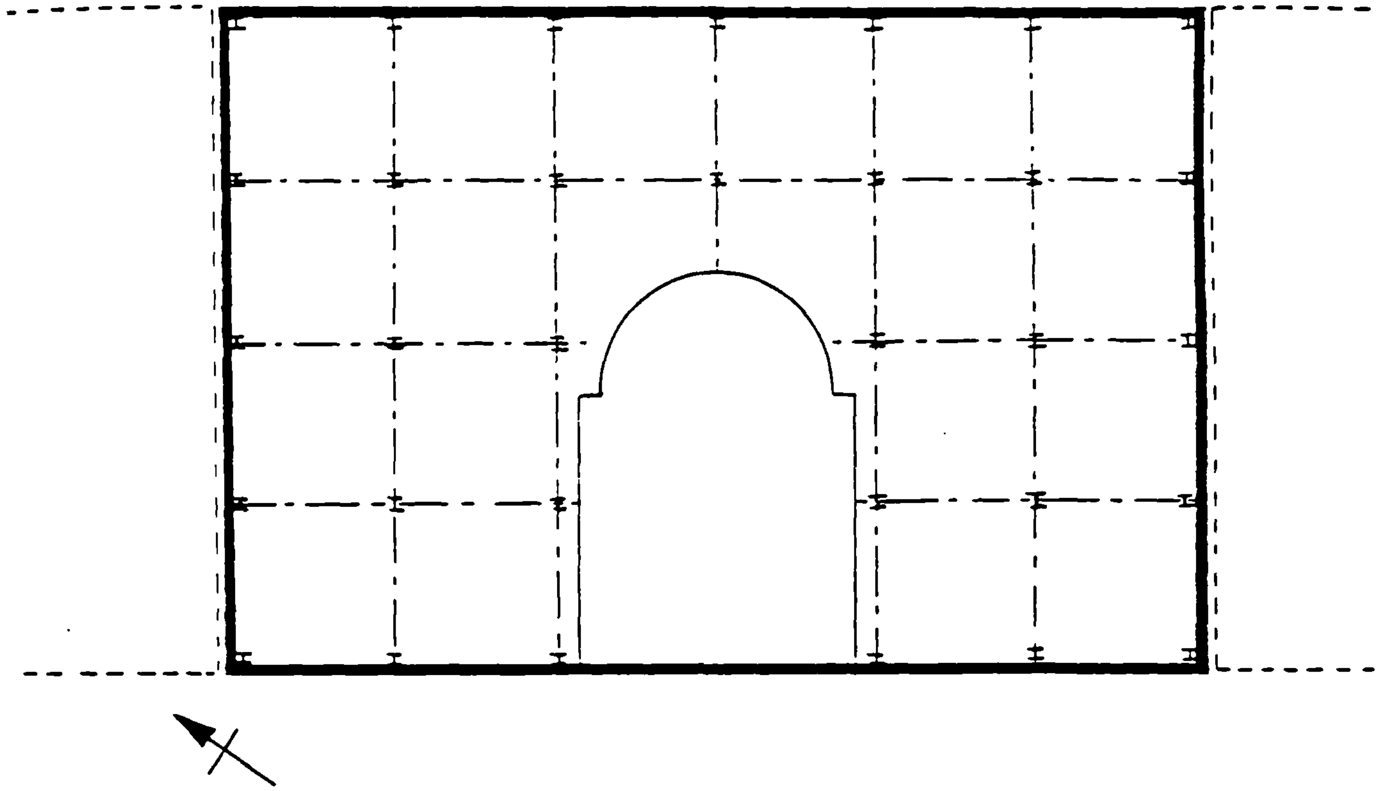
.
.
.

CN 11 note 3 is an exception to this pattern, but see TM 8

Appendix 3

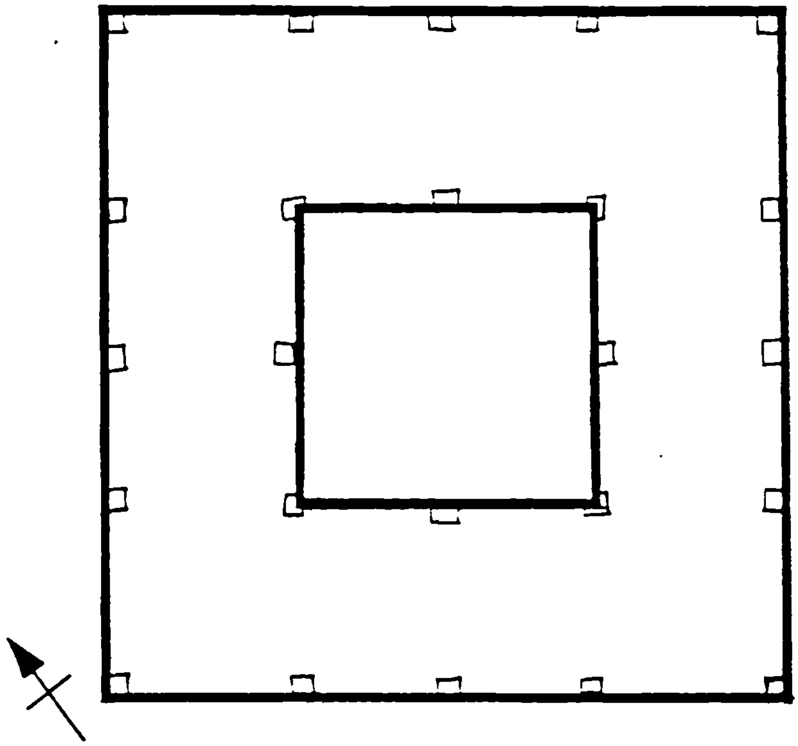
Case Vignettes for the Solution Appraisal Task Study of Chapter 6

Case Vignette 1



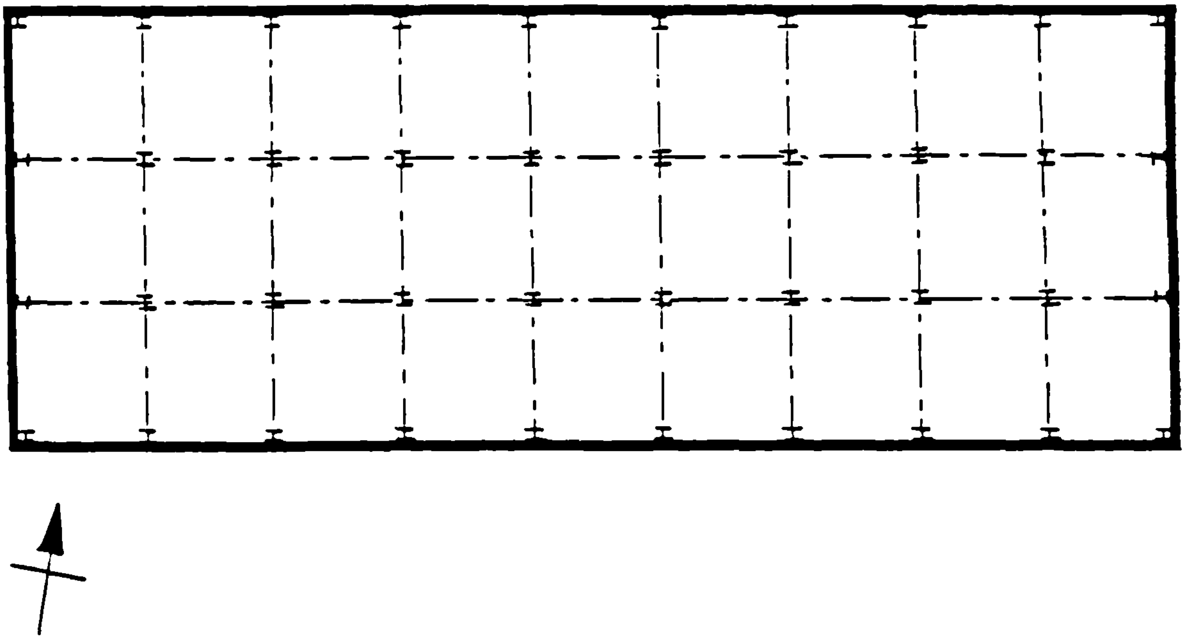
Description	A four storey civil service office of 3,500 sq m in a city centre in line with other existing buildings. Open plan/modular office area 2,000 sq m net.
Structural design	Steel frame with beam depth 400 mm including fire protection. RC floor slab 150 mm thick. 150 mm floor zone required for electrics and datacoms. 2500 mm envisaged floor to ceiling height with slab to slab height negotiable up to 3,700 mm.
Space utilisation	Space required to be open plan but with some modularisation, particularly at the adjoining walls. Power density 20 W/sq m open plan and 10 W/sq m in modular offices. Occupancy density 1 person/9 sq m in open plan areas and 1 person/15 sq m in modular offices.
Long term changes	None specifically anticipated.
Design conditions	External: Winter -5 ^{o}C sat Summer 25 ^{o}C db, 18 ^{o}C wb Internal: Winter 20 ^{o}C Summer < 26 ^{o}C

Case Vignette 2



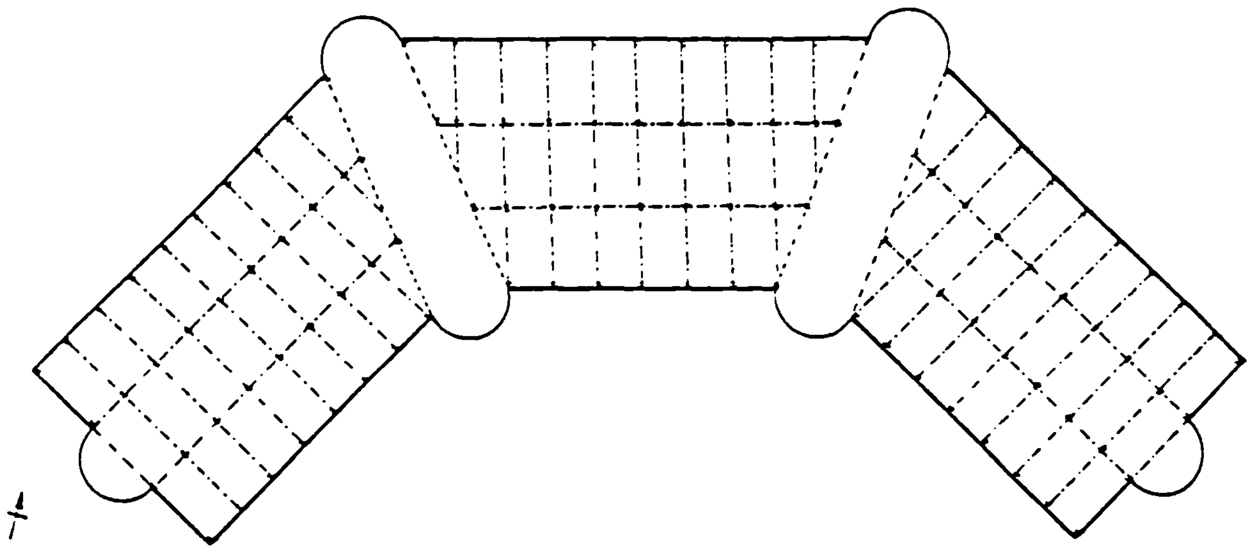
Description	A three storey headquarters building of 4,300 sq m in an edge of town location. Open plan/modular office space of 2,500 sq m net.
Structural design	Concrete waffle slab 600 mm thick. Total slab to slab height negotiable. Electrics and datacoms distributed from perimeter conduits.
Space utilisation	Space required to be open plan, with probable use of low level partitioning. Power density will be uniform at 15 W/sq m. Occupation density 1 person/12 sq m.
Long term changes	Occupancy may increase to 1 person/9 sq m and power density to 20 W/sq m.
Design conditions	External: Winter -5 ^{0}C sat Summer 25 ^{0}C db, 18 ^{0}C wb Internal: Winter 20 ^{0}C Summer N/A

Case Vignette 3



Description	A two storey office building of 2,000 sq m on a research campus for a computer software company. Net open plan/modular office space 1,200 sq m.
Structural design	Steel frame with beam depth 300 mm including fire protection. Floor slab 200 mm. 150 mm floor zone required for electrics and datacoms. Total slab to slab height negotiable.
Space utilisation	Space required to be highly flexible. Initially some areas will be open plan while others will be modular executive offices. There may be a high degree of variation in gains from one area to another due to varied use of micro-computers. Power density in open plan areas 35 W/sq m and in modular offices 10 W/sq m. Occupancy density 1 person/10 sq m in open plan areas and 1 person/20 sq m in modular offices.
Long term changes	None specifically anticipated.
Design conditions	External: Winter -5 ^{0}C sat Summer 25 ^{0}C db, 18 ^{0}C wb Internal: Winter 21 ^{0}C, 40% RH min Summer 22 ^{0}C, 70% RH max

Case Vignette 4

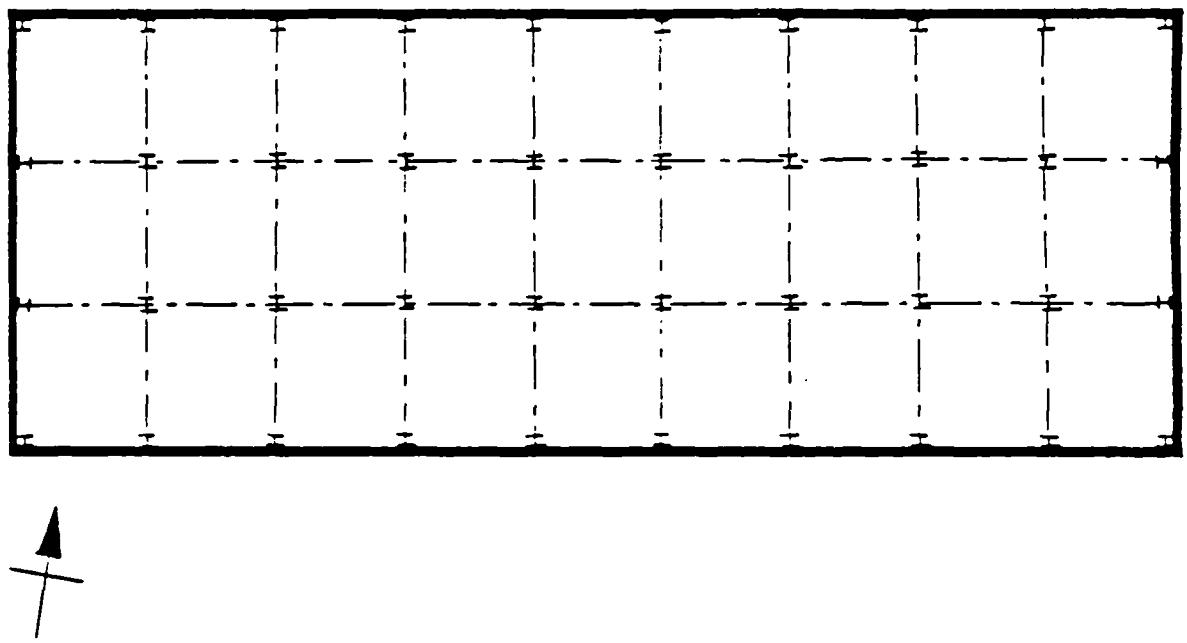


Description	An administration centre on a green field site for a large financial institution. 22,000 sq m in three four storey blocks linked by atria. Open plan/modular office space of 12,000 sq m net.
Structural design	Steel frame with beam depth 450 mm including fire protection. RC floor slab 150 mm. 200 mm floor zone required for electrics and datacoms. 3,000 mm envisaged finished floor to ceiling height with slab to slab height negotiable.
Space division	Generally open plan space with broadly uniform power density of 15 W/sq m and occupancy density of 1 person/10 sq m.
Long term changes	There should be provision for increased power density in some areas of 30 W/sq m.
Design conditions	External: Winter -5 °C sat; Summer 25 °C db, 18 °C wb
	Internal: Winter 21 °C, 40%-70% RH; Summer 22 °C, 40%-70% RH

Appendix 4

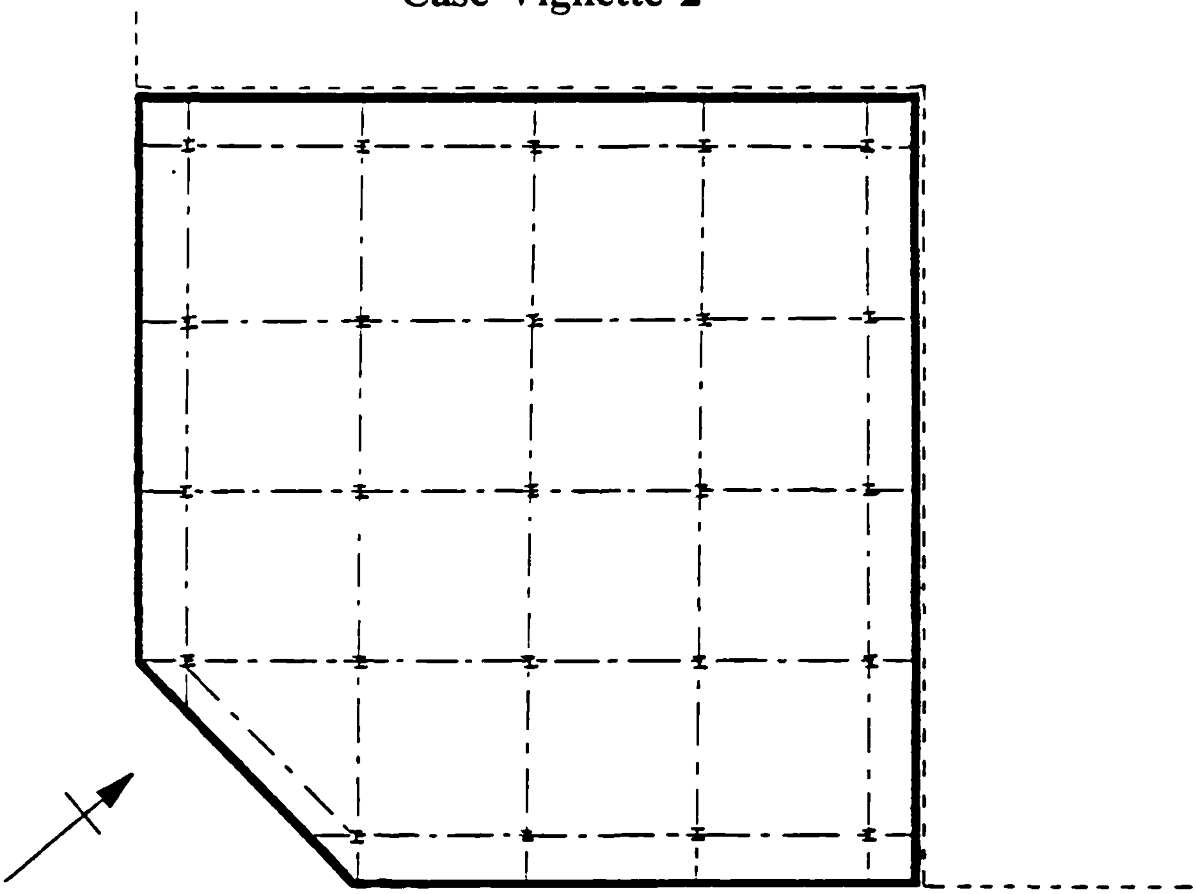
Case Vignettes for the MAVT Validity Study of Chapters 7 and 8

Case Vignette 1



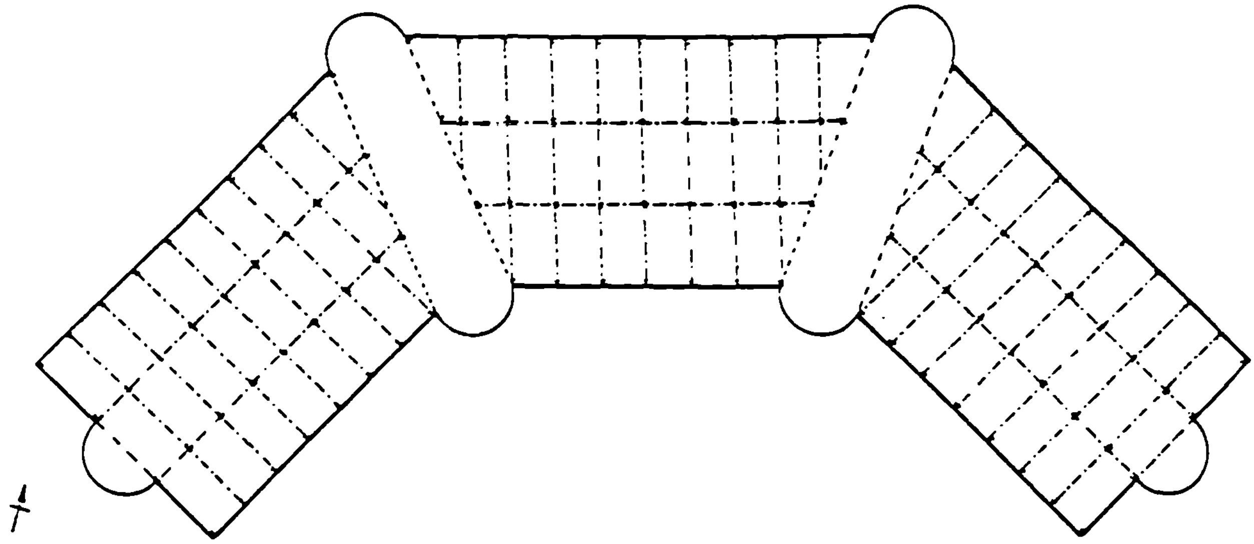
Description	A two storey office building of 2,000 sq m on a research campus for a computer software company.
Structural design	Steel frame with beam depth 300 mm including fire protection. Floor slab 200 mm. 150 mm raised floor for electrics and datacoms. Total slab to slab height negotiable.
Space utilisation	Space required to be highly flexible. Initially some areas will be open plan while others will be modular executive offices. There may be a high degree of variation in gains from one area to another due to varied use of micro-computers. Power density in open plan areas 35 W/sq m and in modular offices 10 W/sq m. Occupancy density 1 person/10 sq m in open plan areas and 1 person/20 sq m in modular offices.
Long term changes	None specifically anticipated.
Maintenance	The company has no maintenance staff of its own and would need to make all maintenance arrangements with a contractor.
Client requests	Humidity control to prevent static build up.

Case Vignette 2



Description	A six storey headquarters building of 6,000 sq m in a city corner gap site. Open plan/modular office space of 4,000 sq m net.
Structural design	Steel frame with beam depth 400mm including fire protection. RC floor slab 150 mm. 150 mm raised floor for datacoms. 2,700 mm envisaged finished floor to ceiling height with slab to slab height negotiable up to 3,900 mm.
Space utilisation	Basic requirement of space to be open plan, but there may be some modularisation. Power density will be similar zone to zone at 15 W/sq m. Occupation density 1 person/10 sq m.
Long term changes	Could be a high degree of modularisation in future, but gains will be generally uniform.
Maintenance	Company has some of its own maintenance staff with experience of all-air and fan coil systems.
Client requests	Concerned about sick building syndrome and require HVAC system to be as clean as possible with good ventilation.

Case Vignette 3

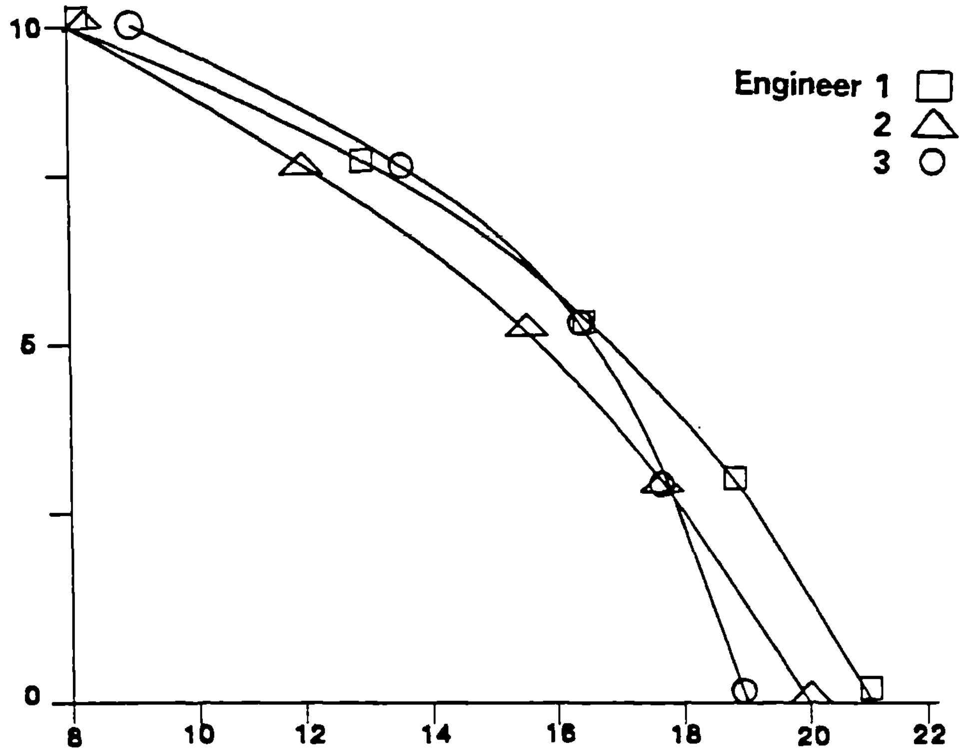


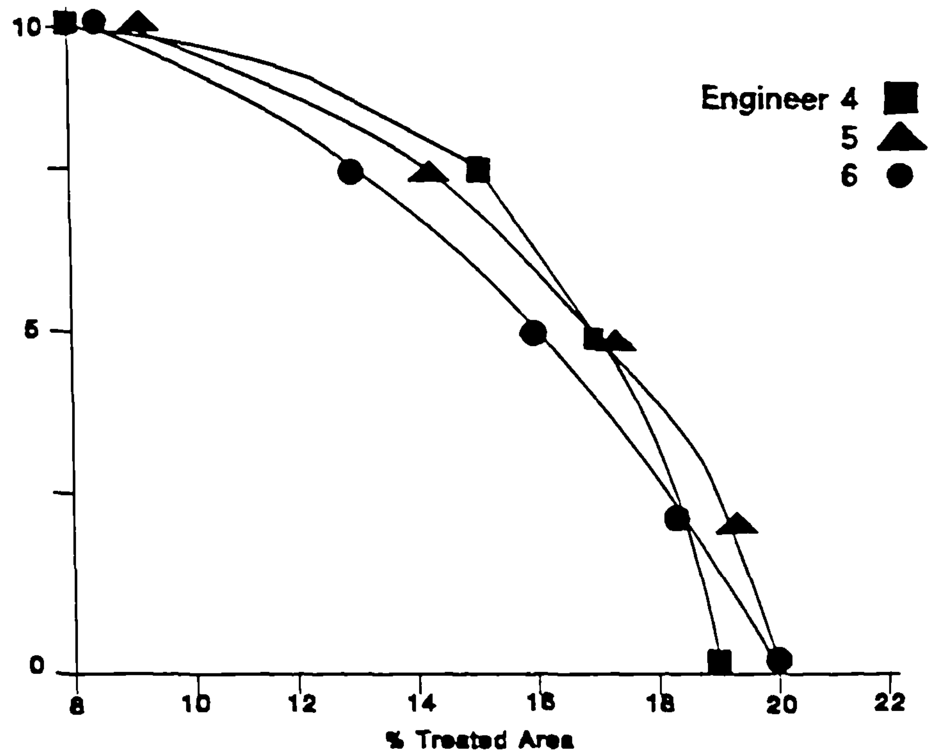
Description	An administration centre on a green field site for a large financial institution. 22,000 sq m in three four storey blocks linked by closed atria. Open plan/modular office space of 12,000 sq m net.
Structural design	Steel frame with beam depth 450 mm including fire protection. RC floor slab 150 mm. 200 mm raised floor for electrics and datacoms. 3,000 mm envisaged finished floor to ceiling height with slab to slab height negotiable.
Space division	Generally open plan space with broadly uniform power density of 15 W/sq m and occupancy density of 1 person/10 sq m.
Long term changes	There should be provision for increased power density in some areas of 30 W/sq m.
Maintenance	The client has highly qualified maintenance staff.
Client requests	The HVAC system should minimise noise and maintain a healthy working environment.

Appendix 5

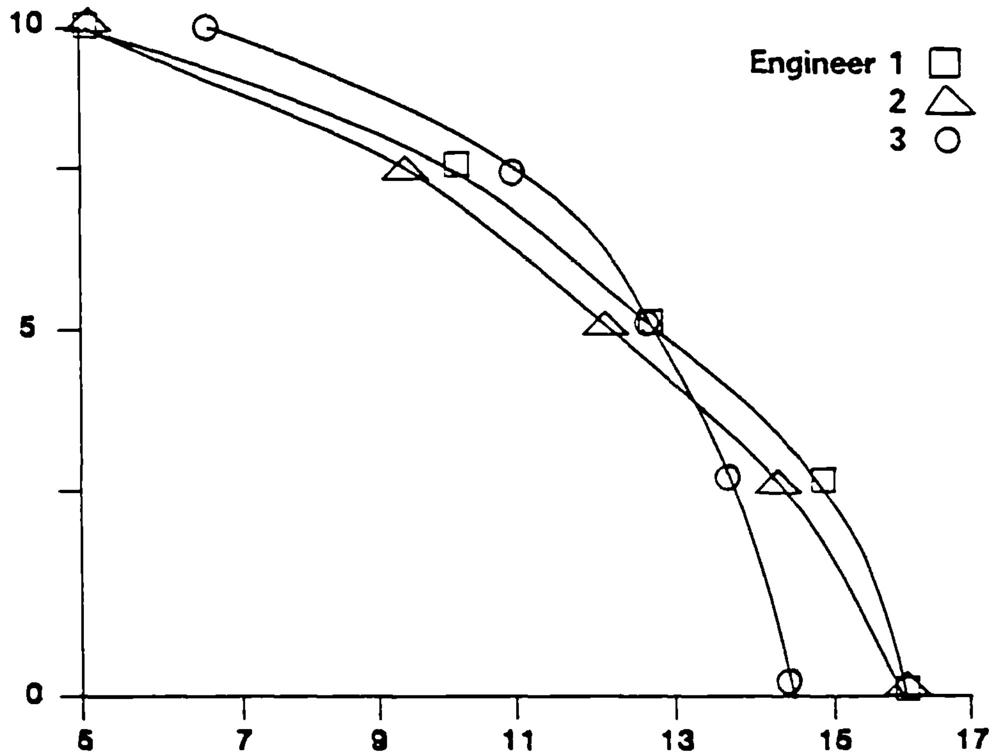
Attribute Value Functions Assessed for the MAVT Validity Study of Chapters 7 and 8

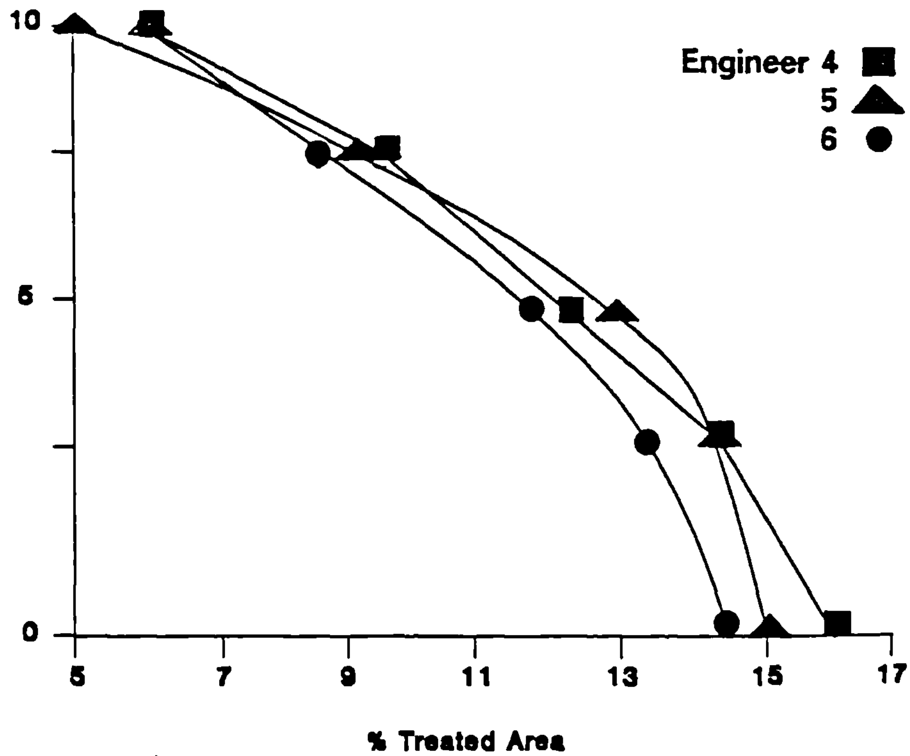
PLANT ROOM AND RISER AREA - CASE 1



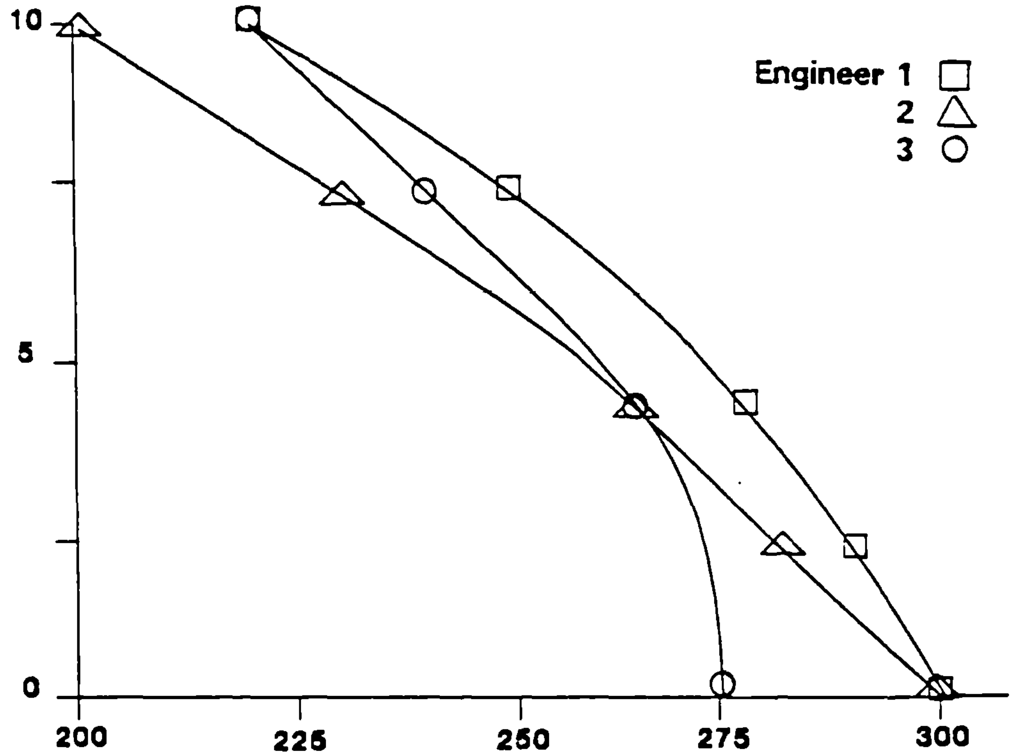


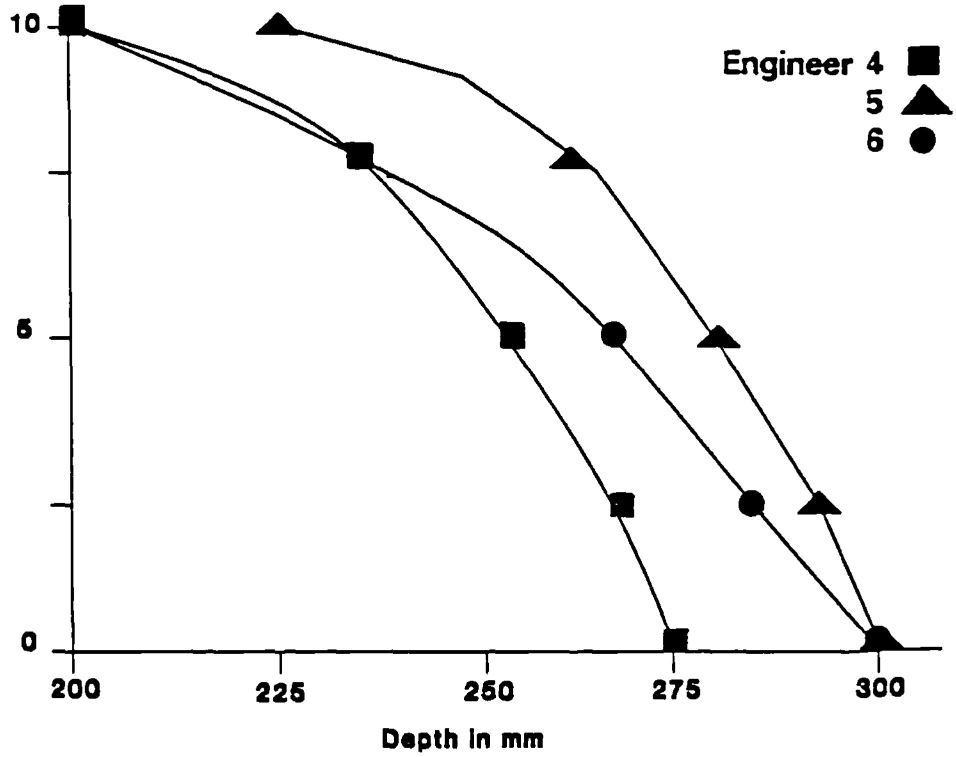
PLANT ROOM AND RISER AREA - CASES 2 & 3



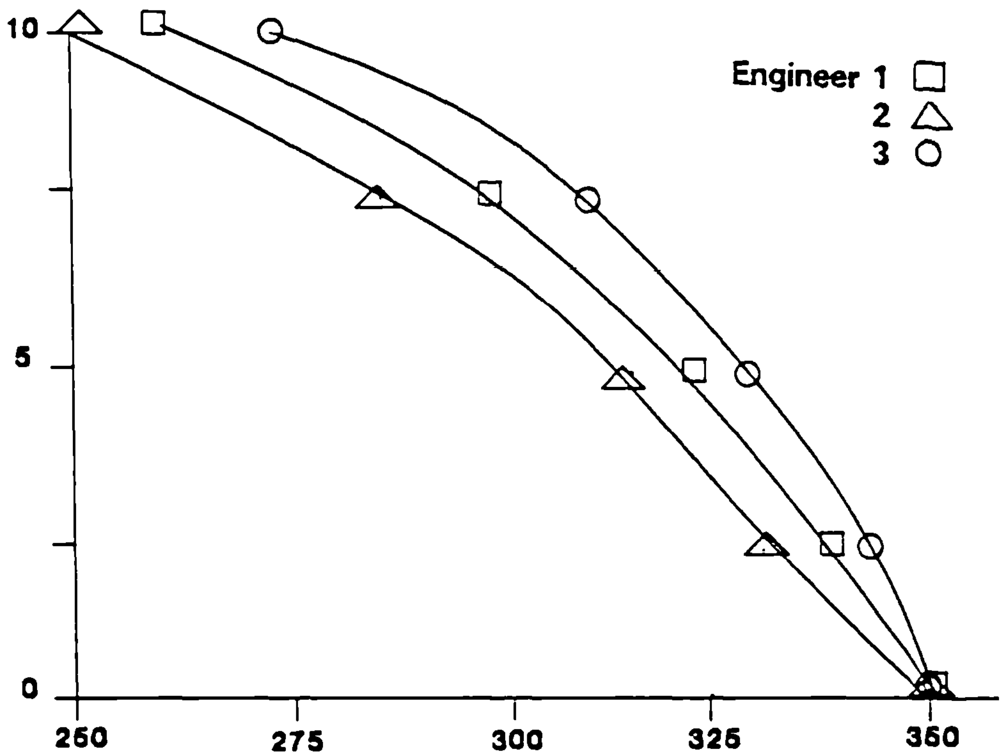


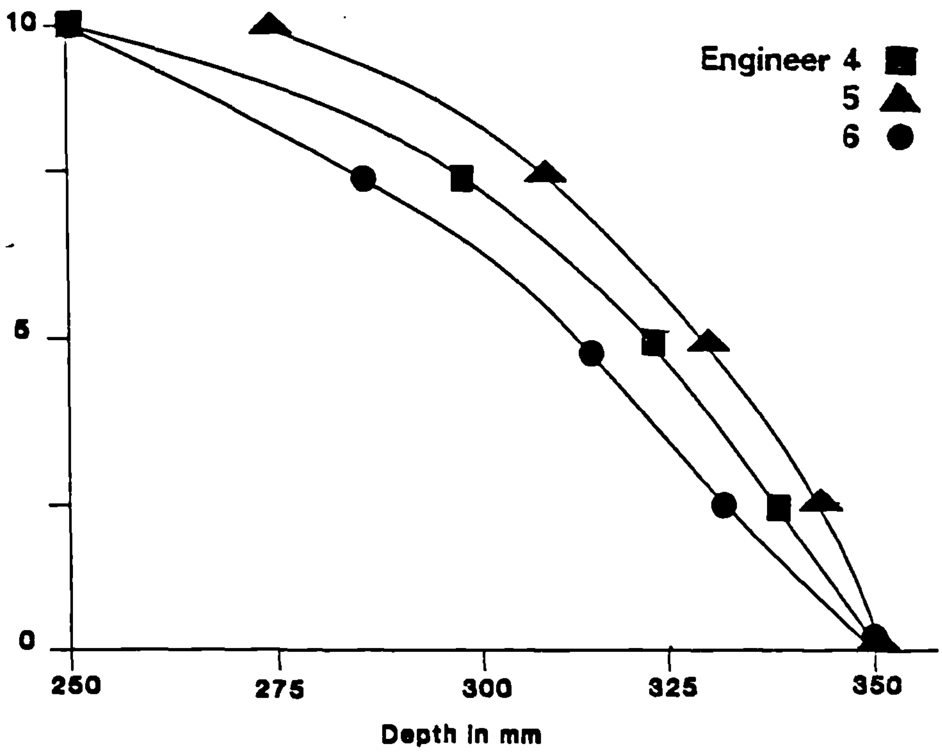
CEILING ZONE DEPTH - CASE 1



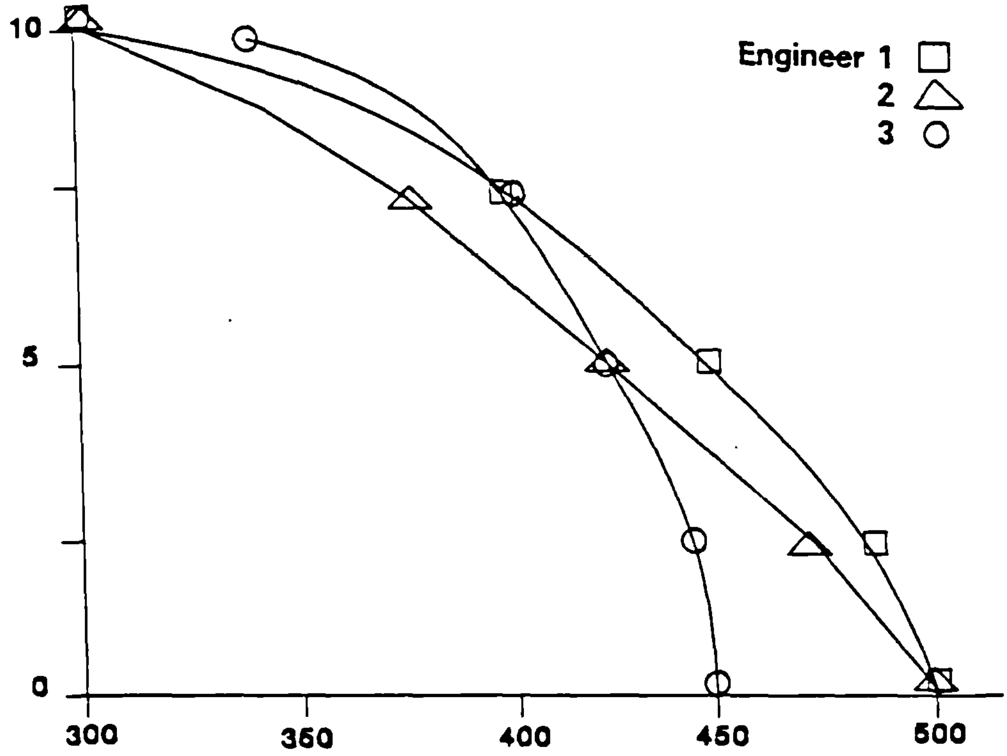


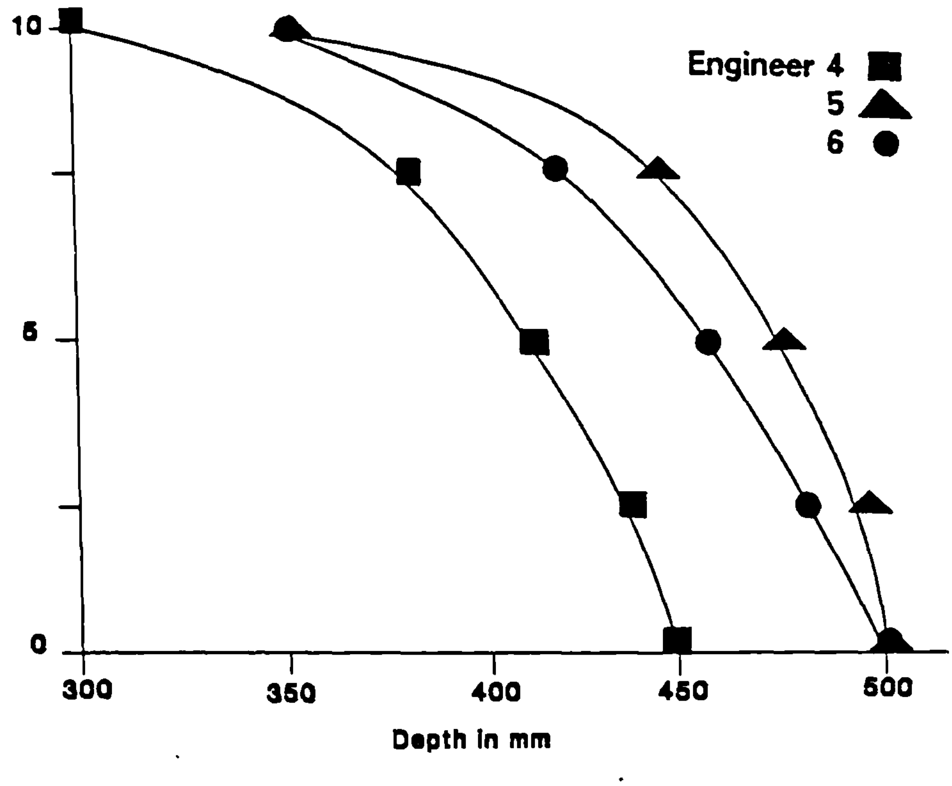
CEILING ZONE DEPTH - CASE 2



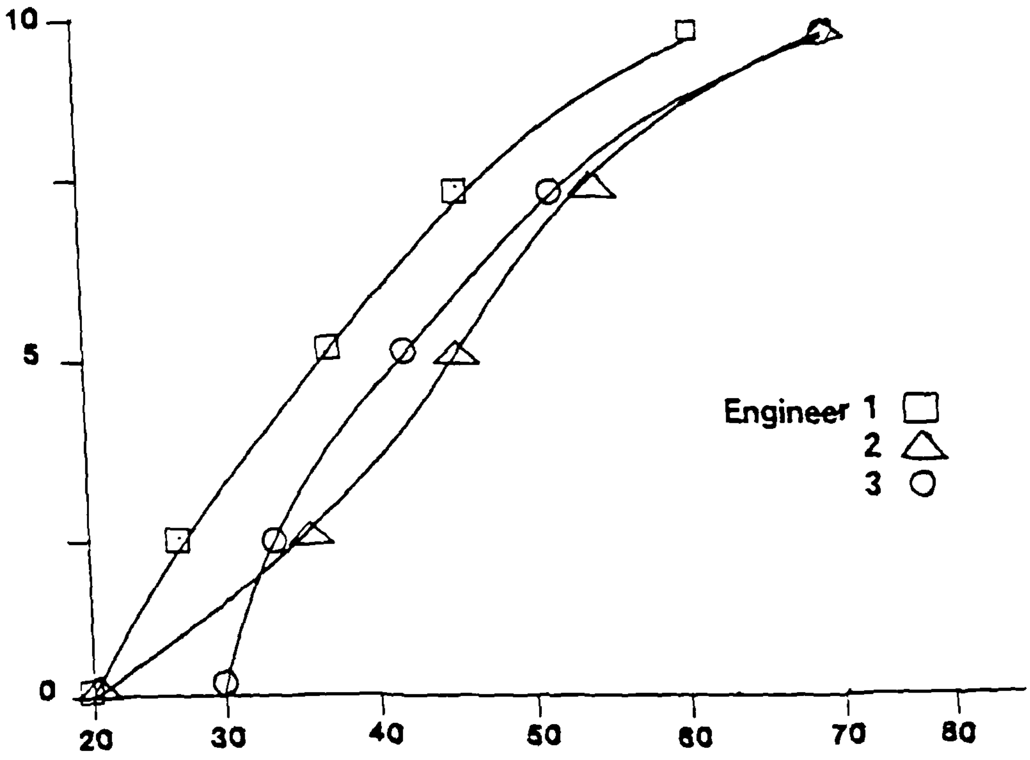


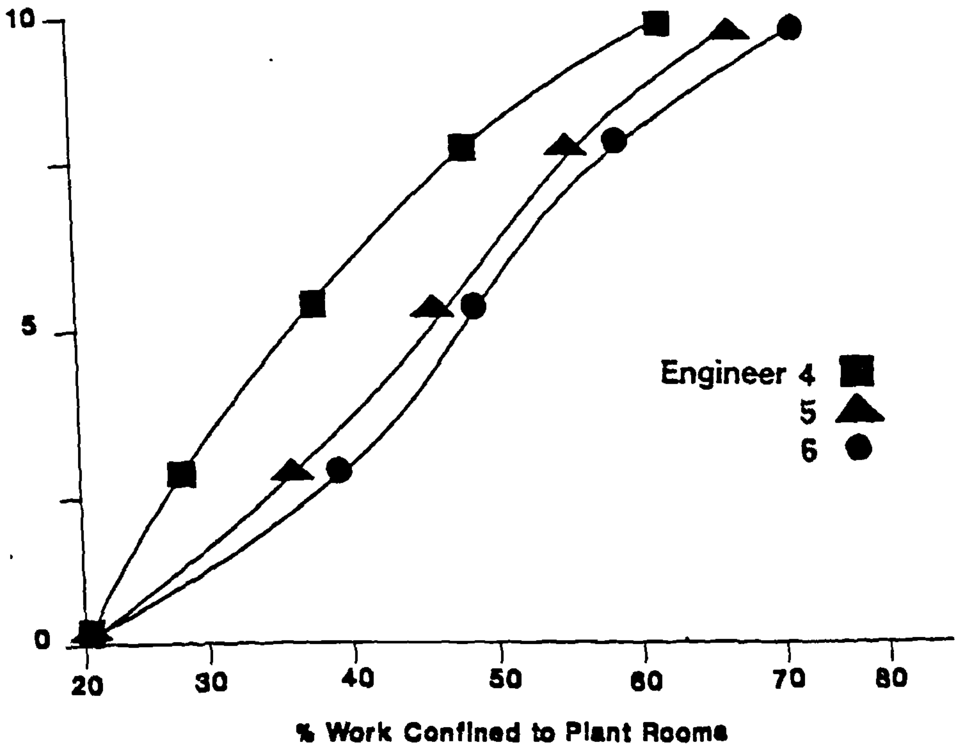
CEILING ZONE DEPTH - CASE 3



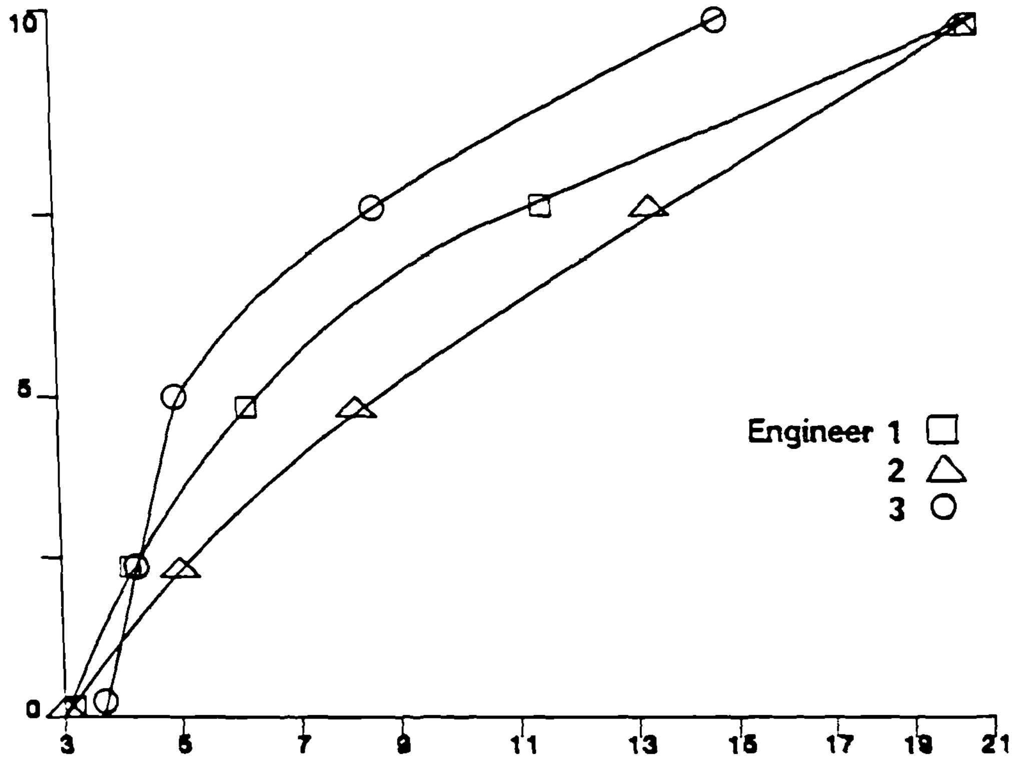


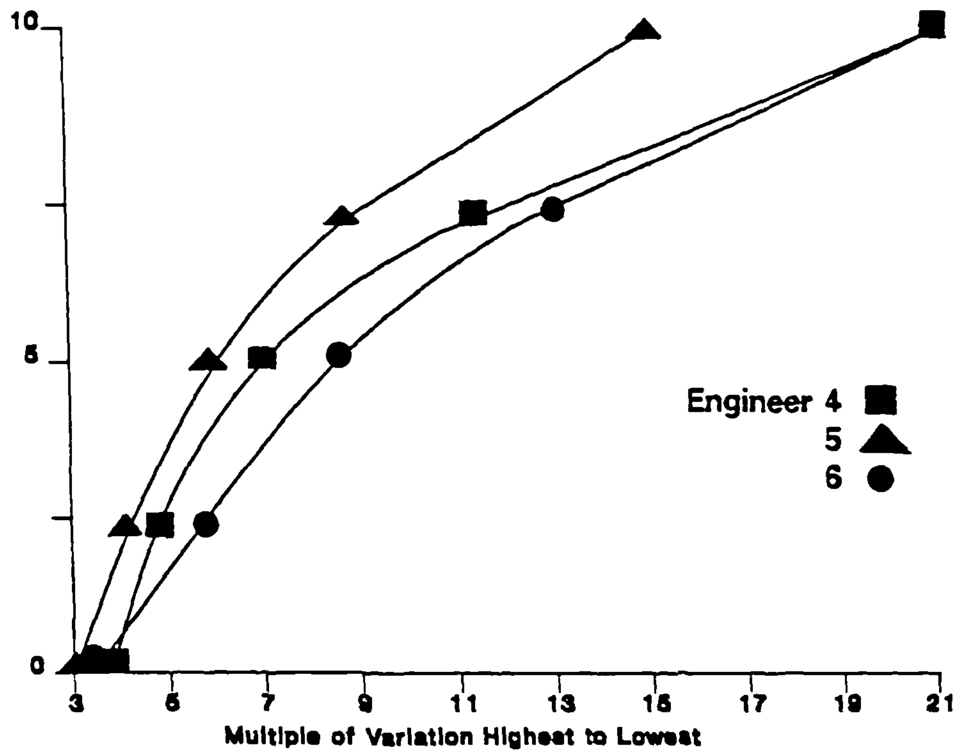
MAINTENANCE NUISANCE - CASES 1, 2 & 3



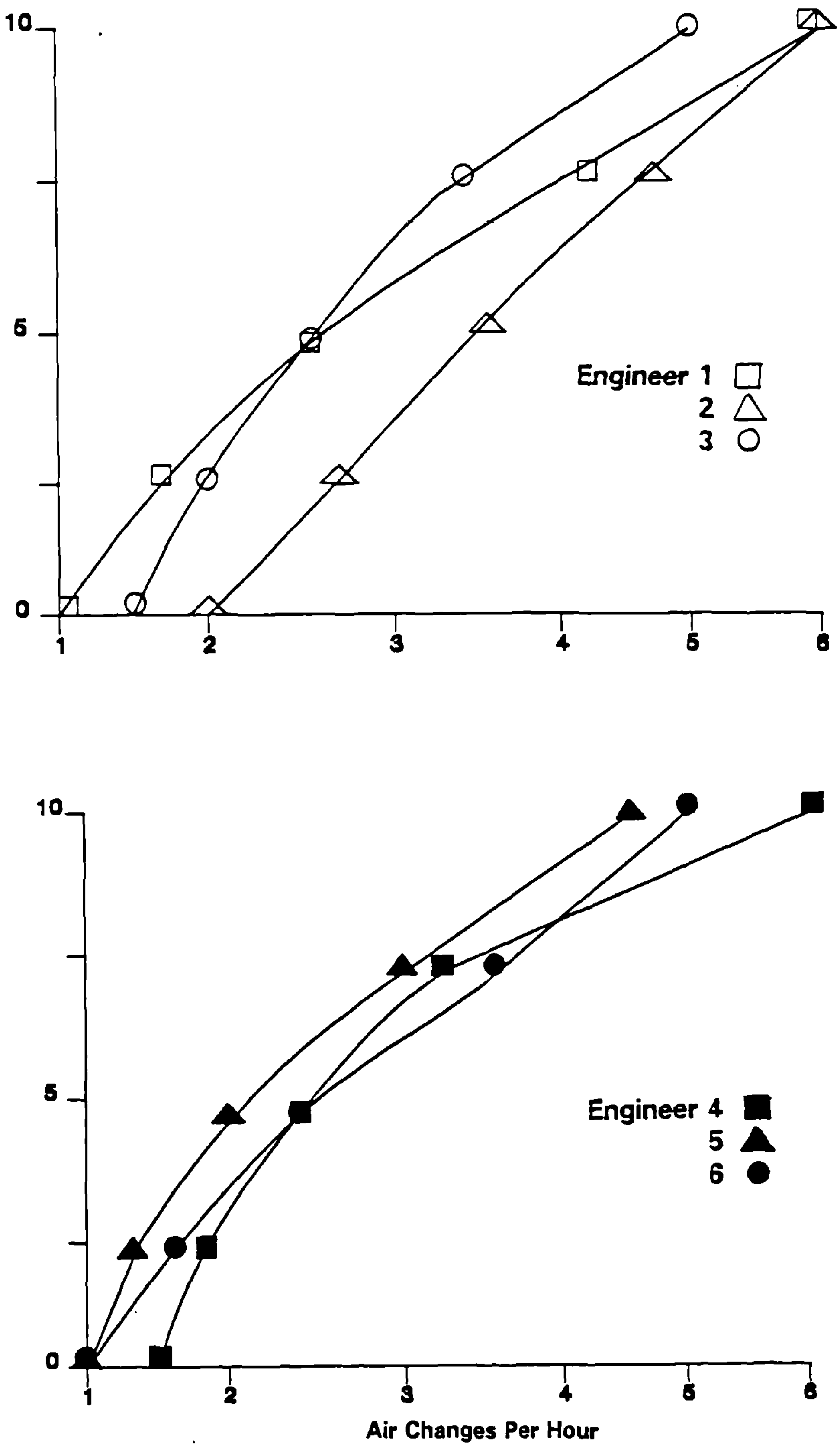


LOAD VARIATIONS HANDLED - CASES 1, 2 & 3

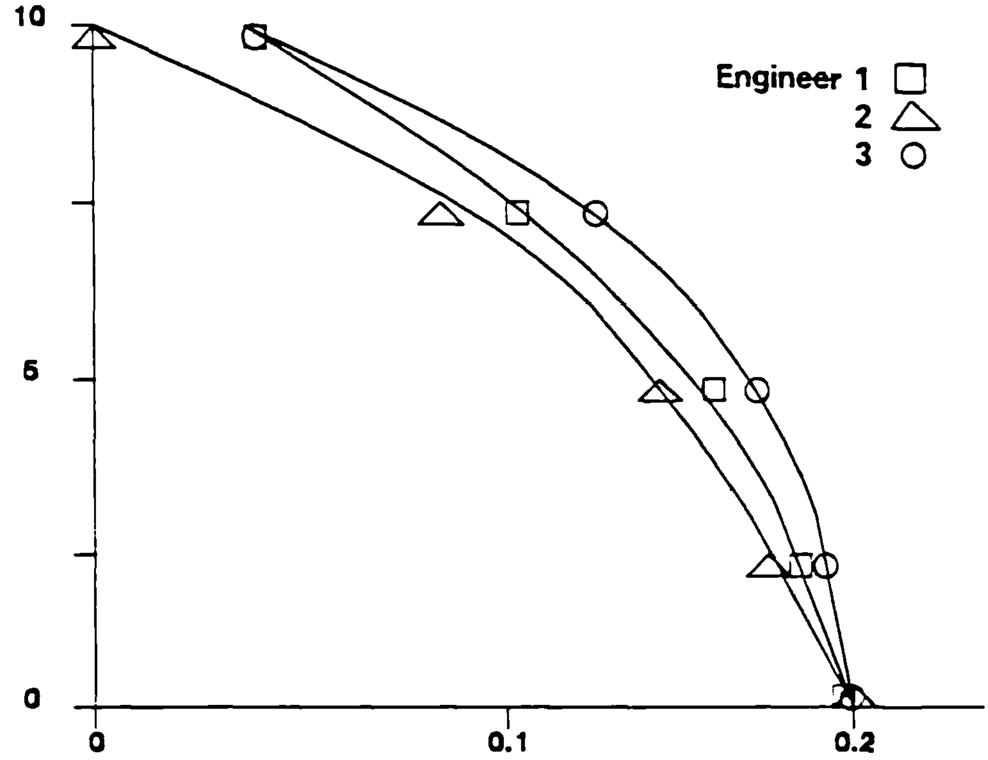


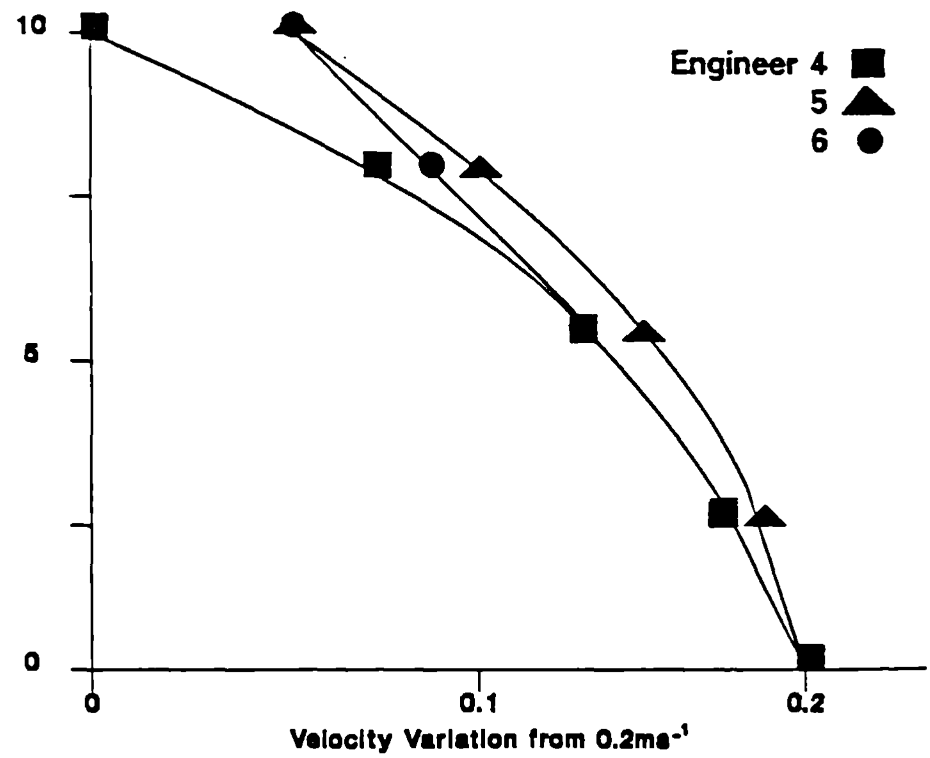


VENTILATION RATES - CASES 1, 2 & 3

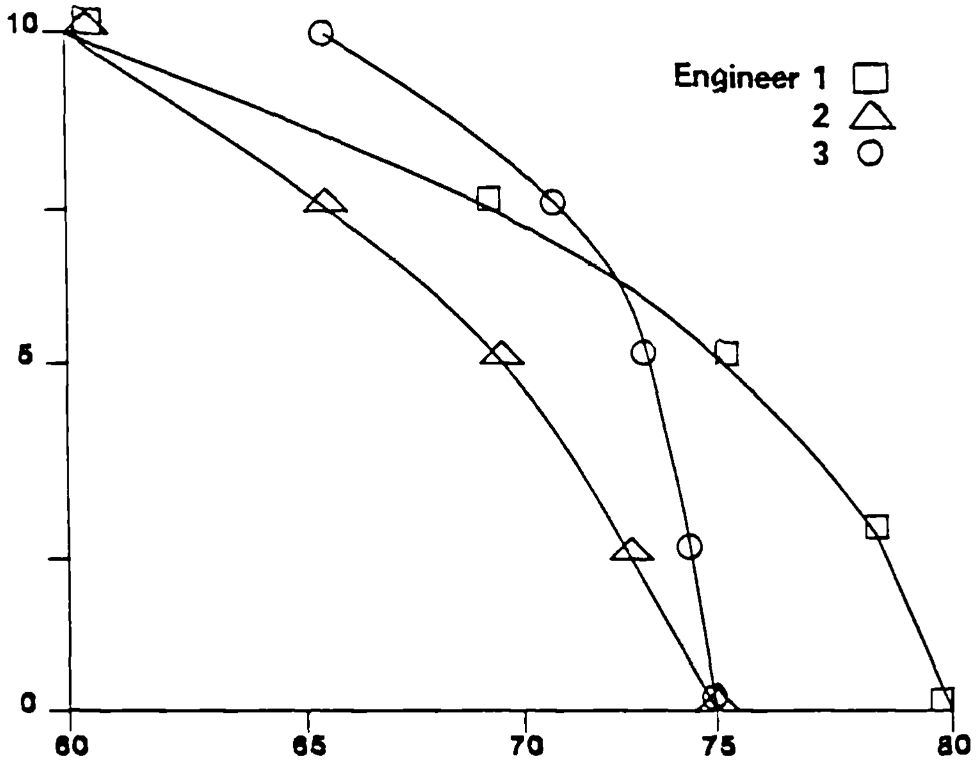


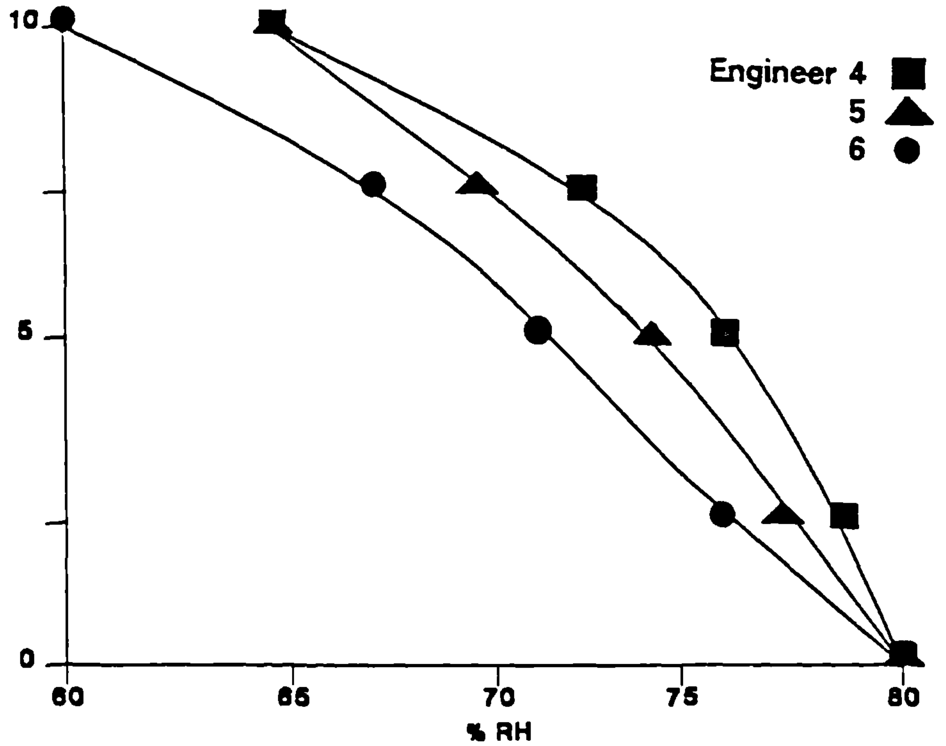
AIR DISTRIBUTION - CASES 1, 2 & 3



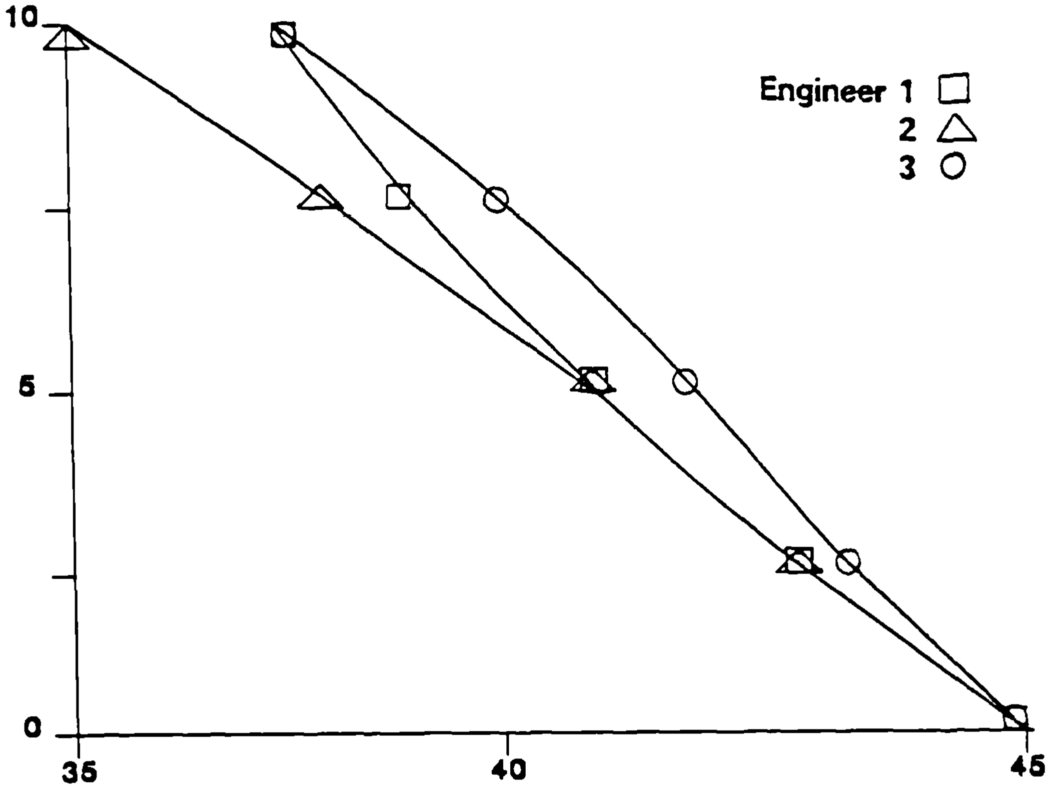


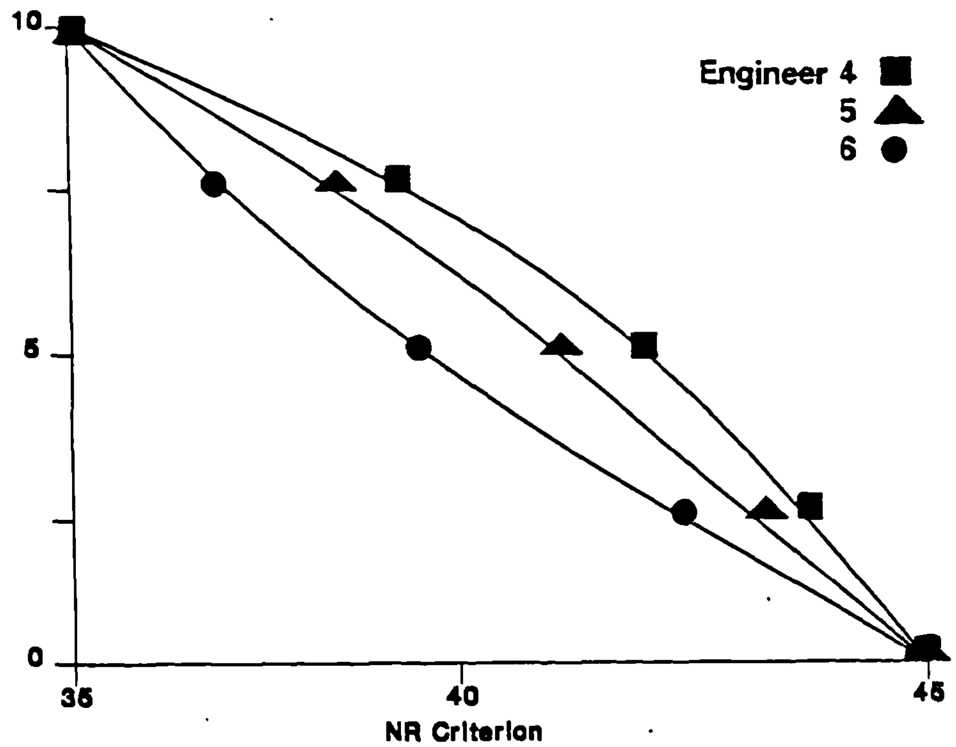
MAXIMUM HUMIDITY - CASES 1, 2 &3



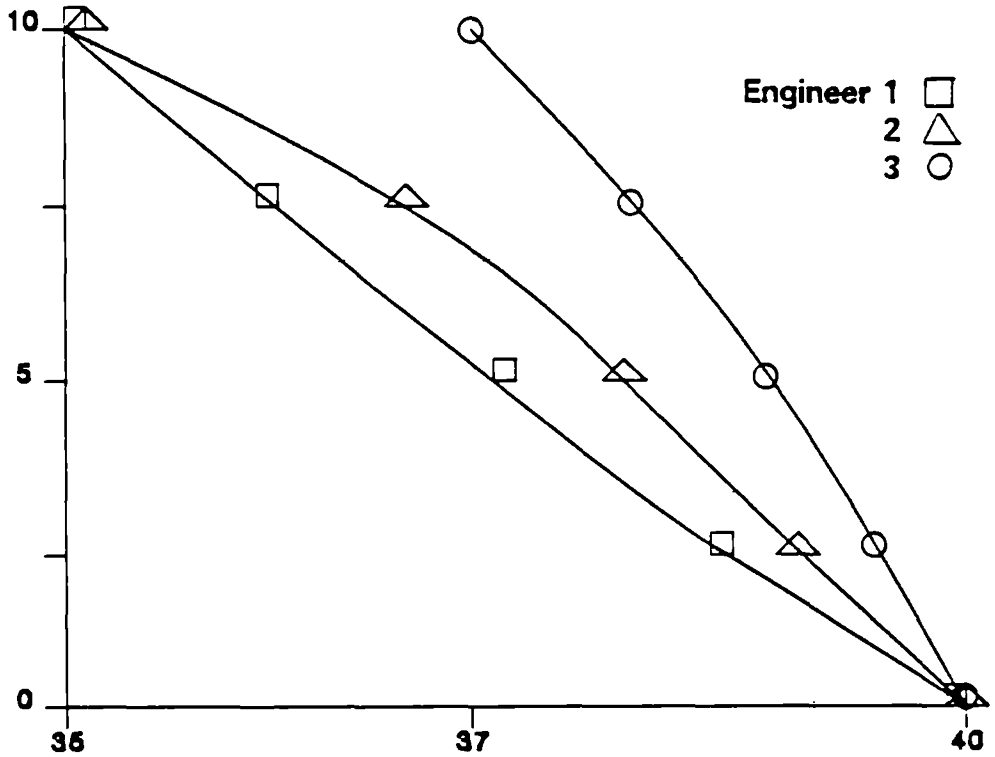


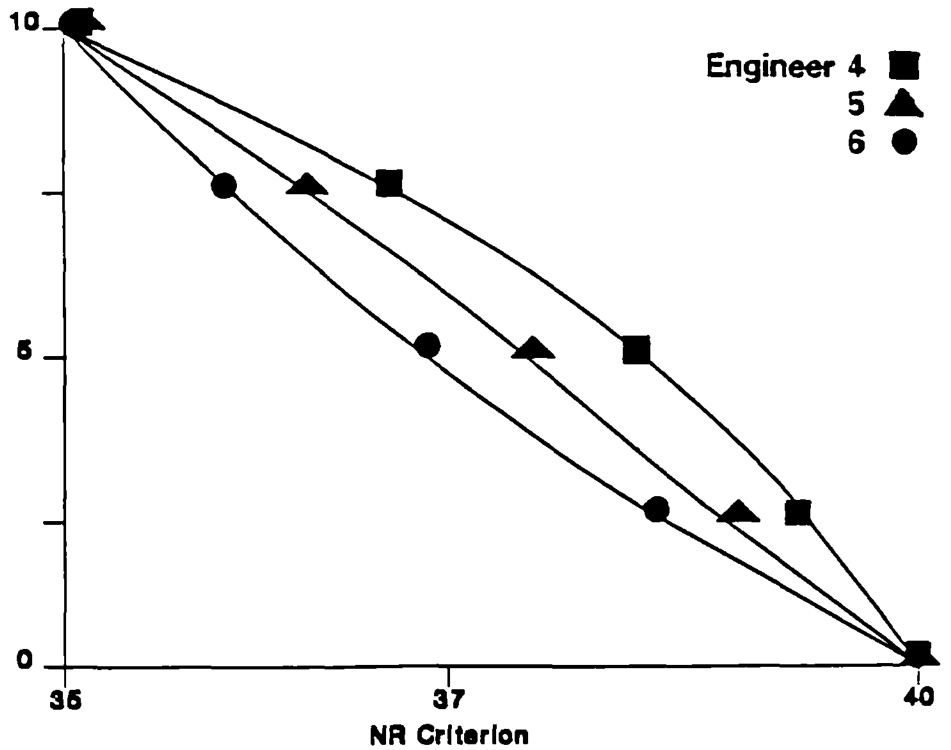
NOISE PRODUCTION - CASES 1 & 2



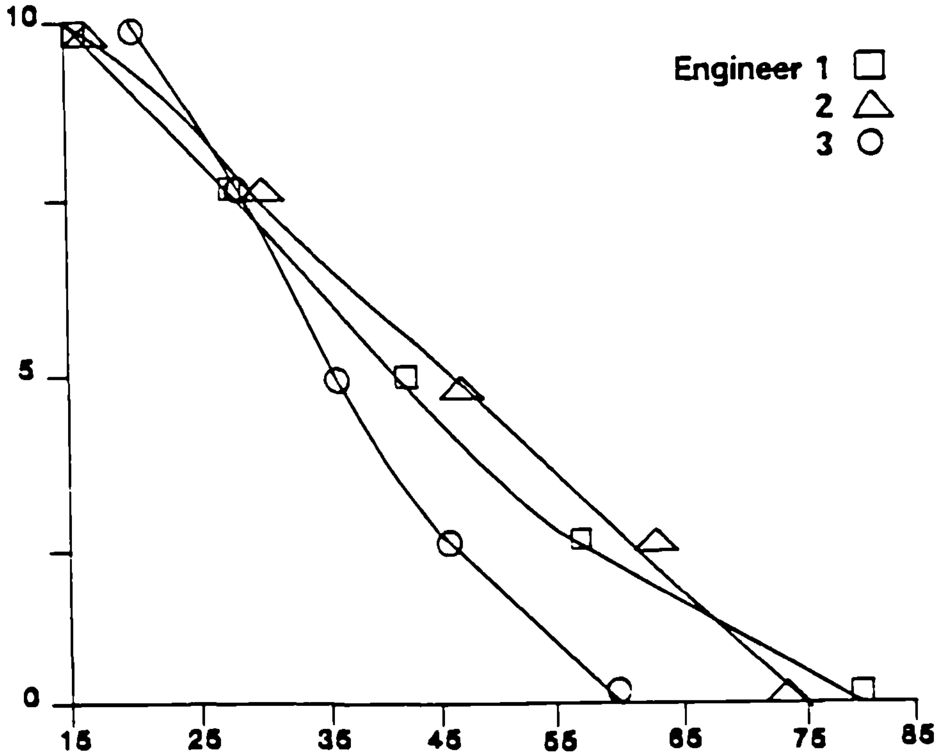


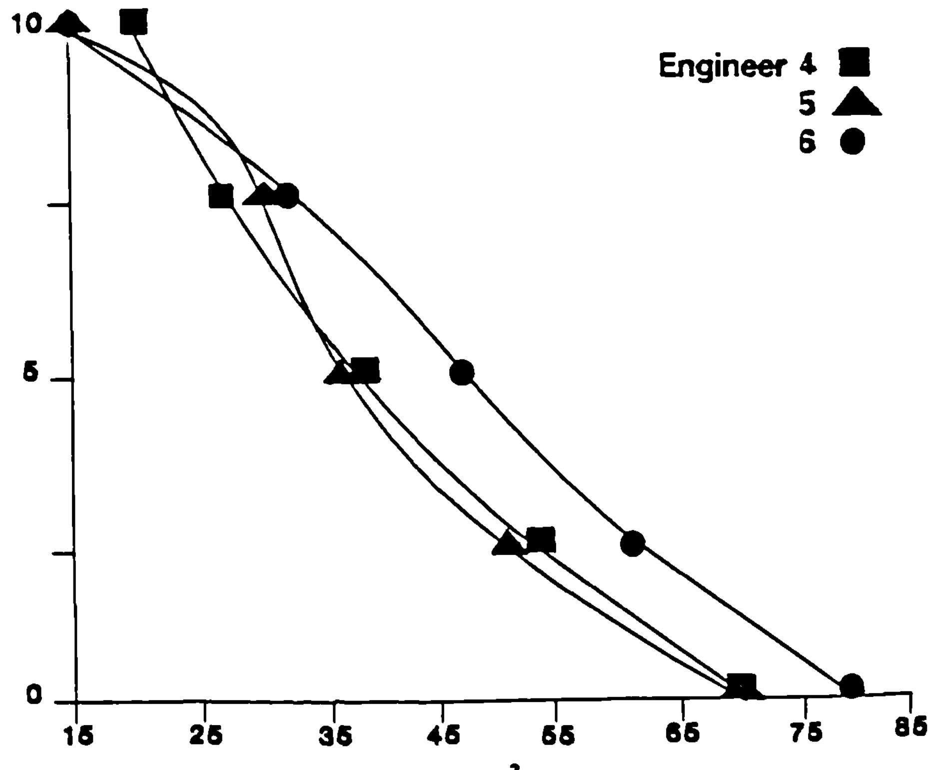
NOISE PRODUCTION - CASE 3



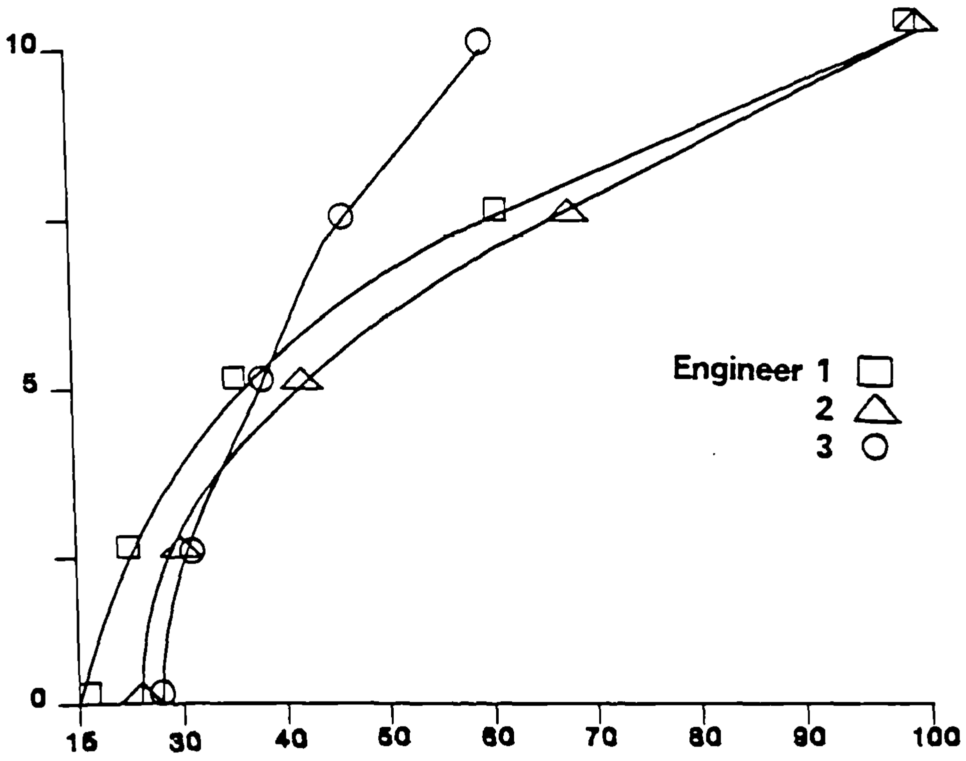


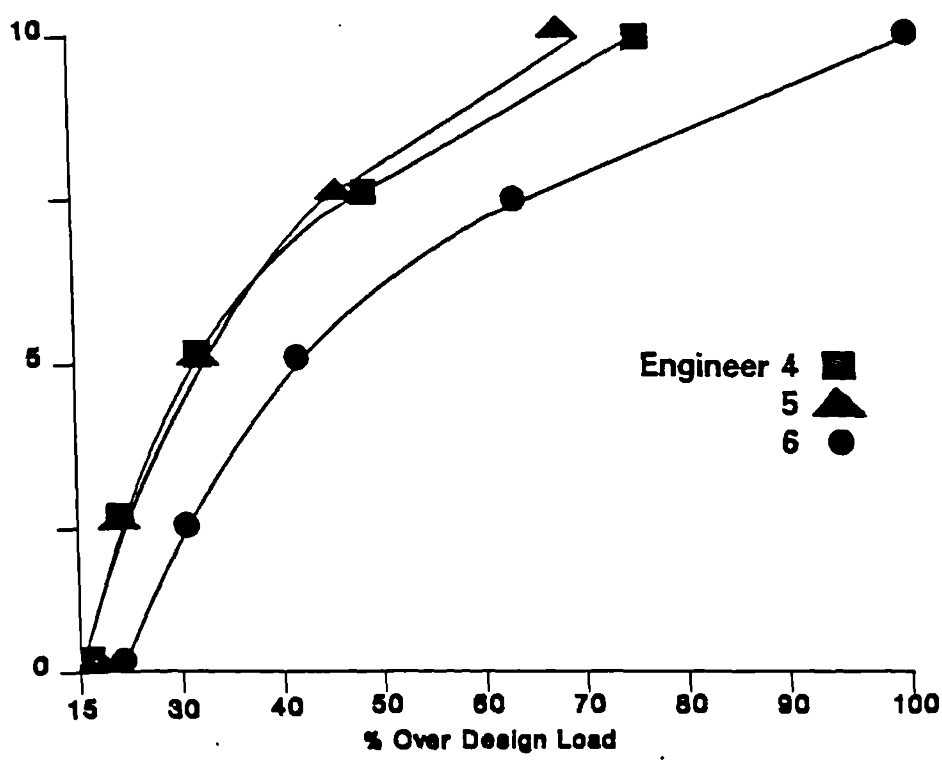
CONTROL ZONE SIZE - CASES 1, 2 &3



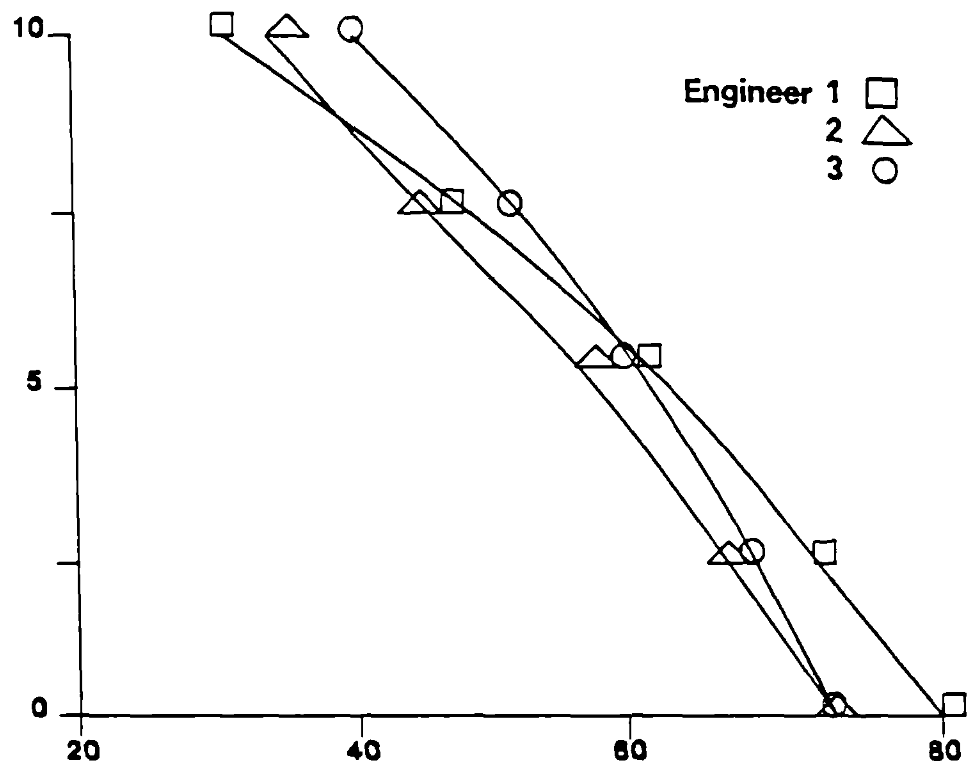


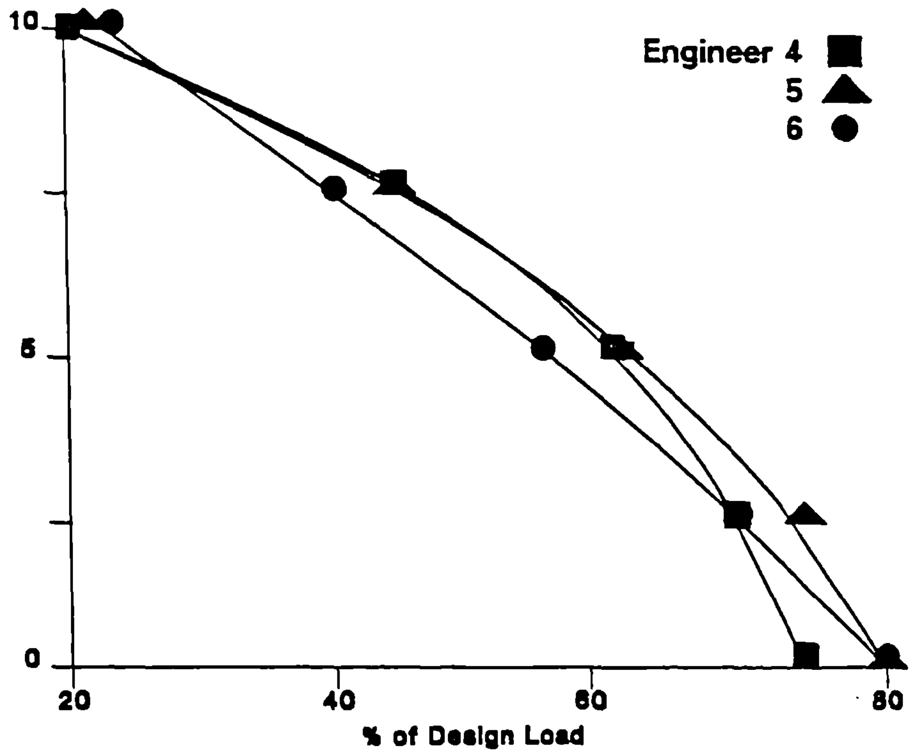
LOAD INCREASE CONTINGENCY - CASES 1, 2 & 3





REDUCED LOAD PERFORMANCE - CASES 1,2 & 3





Appendix 6

Weighted Scoring Tables for the MAVT Validity Study of Chapters 7 and 8

ENGINEER 1

CASE 1							
Attrib	Weight	4PFC	x	VRV	x	RHP	x
1	6	5	30	9	72	7	42
2	1	5	5	10	10	10	10
3	2	4	8	3	6	0	0
4	7	9	63	3	21	1	7
5	22	10	220	8	176	8	176
6	4	6	24	4	16	3	12
7	3	4	12	2	6	0	0
8	1	5	5	5	5	3	3
9	1	7	7	5	5	0	0
10	37	9	333	2	74	6	222
11	7	10	70	7	49	3	21
12	9	9	81	8	72	3	27
		Σ	858	Σ	512	Σ	520

CASE 2									
Attrib	Weight	FATV	x	4PFC	x	VRV	x	RHP	x
1	23	4	92	6	138	9	207	8	184
2	1	3	3	8	8	10	10	10	10
3	16	7	112	4	64	3	48	0	0
4	5	7	35	9	45	3	15	1	5
5	3	6	18	10	30	8	24	8	24
6	13	9	117	6	78	4	52	3	39
7	3	9	27	4	12	2	6	0	0
8	1	8	8	5	5	5	5	3	3
9	1	8	8	7	7	5	5	0	0
10	31	6	186	9	279	2	62	6	186
11	1	0	0	10	10	7	7	3	3
12	2	6	12	9	18	8	16	3	6
		Σ	618	Σ	694	Σ	457	Σ	460

CASE 3							
Attrib	Weight	VAV	x	FATVAV	x	4PFC	x
1	30	2	60	4	120	6	180
2	6	2	12	6	36	8	48
3	40	10	400	7	280	4	160
4	1	7	7	7	7	9	9
5	3	2	6	6	18	10	30
6	8	8	64	9	72	6	48
7	3	8	24	9	18	4	12
8	1	8	8	8	8	5	5
9	1	10	10	5	5	1	1
10	3	5	15	6	18	9	27
11	2	2	4	0	0	10	20
12	2	5	10	6	12	9	18
		Σ	620	Σ	594	Σ	558

CASE 1							
Attrib	Weight	4PFC	x	VRV	x	RHP	x
1	11	4	44	9	99	7	77
2	1	4	4	10	10	10	10
3	1	3	3	3	3	1	1
4	5	8	40	2	10	0	0
5	31	10	310	6	186	6	186
6	4	3	12	1	4	0	0
7	3	3	9	0	0	0	0
8	1	3	3	0	0	0	0
9	1	6	6	6	6	0	0
10	31	9	279	2	62	7	217
11	5	10	50	6	30	1	5
12	6	9	54	5	30	1	6
		Σ	814	Σ	440	Σ	502

CASE 2									
Attrib	Weight	FATV	x	4PFC	x	VRV	x	RHP	x
1	32	3	96	6	192	9	288	8	256
2	1	2	2	6	6	10	10	10	10
3	12	6	72	3	36	3	36	1	12
4	3	8	24	8	24	2	6	0	0
5	2	4	8	10	20	6	12	6	12
6	21	9	189	3	63	1	21	0	0
7	3	9	27	3	9	0	0	0	0
8	1	7	7	3	3	0	0	0	0
9	1	7	7	6	6	6	6	0	0
10	20	7	140	9	180	2	40	7	140
11	2	0	0	10	20	6	12	1	2
12	2	5	10	9	18	5	10	1	2
		Σ	582	Σ	577	Σ	441	Σ	434

CASE 3							
Attrib	Weight	VAV	x	FATVAV	x	4PFC	x
1	37	1	37	3	111	6	222
2	9	2	18	5	45	8	72
3	18	9	162	6	108	3	54
4	1	8	8	8	8	8	8
5	3	1	3	4	12	10	30
6	13	8	104	9	117	3	39
7	4	8	32	9	36	3	12
8	1	7	7	7	7	3	3
9	1	10	10	6	6	0	0
10	3	6	18	7	21	9	27
11	5	1	5	0	0	10	50
12	5	3	15	5	25	9	45
		Σ	419	Σ	496	Σ	562

ENGINEER 3

CASE 1							
Attrib	Weight	4PFC	x	VRV	x	RHP	x
1	5	5	25	10	50	8	40
2	1	0	0	10	10	10	10
3	5	4	20	4	20	0	0
4	11	9	99	2	22	0	0
5	29	10	290	9	261	9	261
6	2	6	12	3	6	2	4
7	2	6	12	3	6	0	0
8	1	5	5	0	0	0	0
9	1	8	8	6	6	0	0
10	36	10	360	1	36	7	252
11	2	10	20	8	16	3	6
12	5	10	50	8	40	0	0
		Σ	901	Σ	473	Σ	573

CASE 2									
Attrib	Weight	FATV	x	4PFC	x	VRV	x	RHP	x
1	26	3	78	7	182	10	260	9	234
2	5	0	0	8	40	10	50	10	50
3	26	7	182	4	104	4	104	0	0
4	5	7	35	9	45	2	10	0	0
5	2	7	14	10	20	9	18	9	18
6	12	10	120	6	72	3	36	2	24
7	3	10	30	6	18	3	9	0	0
8	1	9	9	5	5	0	0	0	0
9	1	9	9	8	8	6	6	0	0
10	19	7	133	10	190	1	19	7	133
11	1	0	0	10	10	8	8	3	3
12	2	7	14	10	20	8	16	0	0
			Σ 624		Σ 714		Σ 536		Σ 462

CASE 3							
Attrib	Weight	VAV	x	FATVAV	x	4PFC	x
1	27	0	0	3	81	7	189
2	7	0	0	6	42	10	70
3	34	9	306	7	238	4	136
4	1	7	7	7	7	9	9
5	3	0	0	7	21	10	30
6	9	9	81	10	90	6	54
7	7	9	63	10	70	6	42
8	1	9	9	9	9	5	5
9	1	10	10	7	7	0	0
10	3	5	15	7	21	10	30
11	5	2	10	0	0	10	50
12	2	5	10	7	14	10	20
		Σ	511	Σ	600	Σ	635

ENGINEER 4

CASE 1							
Attrib	Weight	4PFC	x	VRV	x	RHP	x
1	4	5	20	9	36	8	32
2	1	0	0	10	10	10	10
3	2	4	8	3	6	0	0
4	10	9	90	2	20	0	0
5	25	10	250	7	175	7	175
6	5	6	30	3	15	2	10
7	3	3	9	2	6	0	0
8	1	6	6	6	6	2	2
9	1	7	7	5	5	0	0
10	35	10	350	1	35	6	210
11	2	10	20	7	14	4	8
12	11	9	99	7	77	2	22
		Σ	889	Σ	405	Σ	469

CASE 2									
Attrib	Weight	FATV	x	4PFC	x	VRV	x	RHP	x
1	15	4	60	7	105	9	135	8	120
2	1	0	0	7	7	10	10	10	10
3	33	8	264	4	132	3	99	0	0
4	5	6	30	9	45	2	10	0	0
5	3	6	18	10	30	7	21	7	21
6	19	9	171	6	114	3	57	2	38
7	3	9	27	3	9	2	6	0	0
8	1	9	9	6	6	6	6	2	2
9	1	9	9	7	7	5	5	0	0
10	14	6	84	10	140	1	14	6	84
11	2	0	0	10	20	7	14	4	8
12	3	6	18	9	27	7	21	2	6
		Σ	690	Σ	642	Σ	398	Σ	289

CASE 3							
Attribs	Weight	VAV	x	FATVAV	x	4PFC	x
1	18	0	0	4	21	7	126
2	1	0	0	5	5	9	9
3	36	10	360	8	288	4	144
4	6	6	36	6	36	9	54
5	3	0	0	6	18	10	30
6	14	9	126	9	126	6	84
7	9	9	81	9	81	3	27
8	1	9	9	9	9	6	6
9	1	10	10	7	7	2	2
10	3	5	15	6	18	10	30
11	6	2	12	0	0	10	60
12	2	4	8	6	12	9	18
		Σ	657	Σ	672	Σ	590

ENGINEER 5

CASE 1							
Attrib	Weight	4PFC	x	VRV	x	RHP	x
1	8	5	40	9	72	7	56
2	1	5	5	10	10	10	10
3	3	3	9	2	6	1	3
4	8	9	72	3	24	0	0
5	26	10	260	8	208	8	208
6	3	6	18	4	12	4	12
7	2	5	10	2	4	0	0
8	1	5	5	5	5	3	3
9	1	7	7	5	5	0	0
10	36	10	360	1	36	6	216
11	4	10	40	8	32	4	16
12	7	9	63	6	42	2	14
		Σ	889	Σ	456	Σ	538

CASE 2									
Attrib	Weight	FATV	x	4PFC	x	VRV	x	RHP	x
1	25	4	100	7	175	9	225	8	200
2	3	3	9	7	21	10	30	10	30
3	20	8	160	3	60	2	40	1	20
4	4	6	24	9	36	3	12	0	0
5	2	6	12	10	20	8	16	8	16
6	15	10	150	6	90	4	60	4	60
7	4	10	40	5	20	2	8	0	0
8	1	8	8	5	5	5	5	3	3
9	1	8	8	7	7	5	5	0	0
10	20	5	100	10	200	1	20	6	120
11	2	1	2	10	20	8	16	4	8
12	3	6	18	9	27	6	18	2	6
		Σ	631	Σ	681	Σ	455	Σ	463

CASE 3							
Attrib	Weight	VAV	x	FATVAV	x	4PFC	x
1	31	2	62	4	124	7	217
2	7	3	21	7	49	10	70
3	31	10	310	8	248	3	93
4	3	6	18	6	18	9	27
5	4	1	4	6	24	10	40
6	9	10	90	10	90	6	54
7	5	8	40	10	50	5	25
8	1	8	8	8	8	5	5
9	1	10	10	6	6	0	0
10	2	5	10	5	10	10	20
11	4	2	8	1	4	10	40
12	2	4	8	6	12	9	18
		Σ	589	Σ	643	Σ	609

ENGINEER 6

CASE 1							
Attrib	Weight	4PFC	x	VRV	x	RHP	x
1	9	4	36	9	81	6	54
2	1	5	5	10	10	10	10
3	3	2	6	2	6	0	0
4	10	10	100	3	30	1	10
5	23	10	230	7	161	6	138
6	3	5	15	3	9	2	6
7	2	4	8	2	4	0	0
8	1	3	3	2	2	2	2
9	1	5	5	3	3	0	0
10	33	9	297	2	66	6	198
11	6	10	60	6	36	2	12
12	8	9	72	7	56	2	16
		Σ	837	Σ	464	Σ	446

CASE 2									
Attrib	Weight	FATV	x	4PFC	x	VRV	x	RHP	x
1	22	3	66	6	132	9	198	8	176
2	2	2	4	6	12	10	20	10	20
3	22	8	176	2	44	2	44	0	0
4	4	7	28	10	40	3	12	1	4
5	3	4	12	10	30	7	21	6	18
6	20	10	200	5	100	3	60	2	40
7	4	10	40	4	16	2	8	0	0
8	1	7	7	3	3	2	2	2	2
9	1	7	7	5	5	3	3	0	0
10	18	6	108	9	162	2	36	6	108
11	1	0	0	10	10	6	6	2	2
12	2	5	10	9	18	7	14	2	4
		Σ	658	Σ	572	Σ	424	Σ	374

CASE 3							
Attrib	Weight	VAV	x	FATVAV	x	4PFC	x
1	20	2	40	3	60	6	120
2	5	2	10	6	30	10	50
3	47	9	423	8	376	2	94
4	1	8	8	7	7	10	10
5	3	0	0	4	12	10	30
6	10	9	90	10	100	5	50
7	5	9	45	10	50	4	20
8	1	7	7	7	7	3	3
9	1	10	10	5	5	0	0
10	3	6	18	6	18	9	27
11	2	1	2	0	0	10	20
12	2.	3	6	5	10	9	18
		Σ	659	Σ	675	Σ	442