Royal College of Art

Department of Design Research
Professor Bruce Archer CBE DrRCA CEng MIMechE MIED
Head of Department Richard Langdon MA MDes(RCA)

Kensington Gore
London SW7 2EU

Telephone 01-584 5020

Thesis submission for the degree of PhD

Applicant: Sebastian Galileo Lera BA BArch RIBA

Supervisor: Dr George Mallen BSc PhD MBCA

Designers' Values and the Evaluation of Designs

Sebastian G. Lera May 1980

CONTENTS

LIST OF TABLES AND FIGURES

ACKNOWLEDGEMENTS

This thesis is a product of the rich exchange of ideas experienced in a research environment. I am grateful to all members of the Department of Design Research for providing such an environment. I am also grateful to the Science Research Council for the financial support which made this study possible.

Many individuals contributed directly in a number of ways; I should like to thank the following in particular: Professor Bruce Archer, Dr. Alan Brookes, Peter Coles, Janet Daley, Pierre Goumain, Richard Langdon, Professor John Miller, Professor George L. Smith, Dr. Brian Thompson, Christopher Woodward and David Yeomans.

In all, fifty subjects took part in the experiments and I express my gratitude to them, especially to those who gave up their own free time to take part.

My most essential thanks are due to my supervisor, Dr. George Mallen. He was responsible for recommending many of the literature sources, and provided ideas, motivation and much encouragement for the present study.

Finally I must thank Jennifer-Anne Macmillan for the considerable effort she put into typing this thesis.

DECLARATION

The computer program used to calculate eigenvectors and eigenvalues from a reciprocal matrix was written by Dr. George Mallen. This program, SAATY, is shown in appendix 4.9. The present author embodied the SAATY program into his own program to calculate additive utility model indices, as shown in appendix 4.7. This apart, the thesis is the sole work of the author.

SUMMARY

Given the same brief, different designers propose different designs; this is the basis of architectural competitions and may be seen in any design studio.

This thesis proposes a tentative theory that design alternatives are the result of each designer's subjective structuring of the design problem, of the way he exercises value judgements and of his perception of priorities among competing objectives. The theory also implies that in the evaluation of alternatives, preference will be expressed for those designs which reflect the evaluator's priorities.

The theory is tested experimentally. In each of five intensive design exercises, architectural designers prepare sketch designs. In conjunction with the design process they perform a judgement analysis exercise, recently devised by Thomas Saaty, intended to elicit their subjective structuring of the problem. The sketch designs, having been redrawn, are then evaluated by the designers. The evaluation makes use, again, of Saaty's technique to elicit each designer's partial judgements of the designs; an additive utility model enables these partial judgements to be combined into an index, expressing overall preferences.

Data from the experiments enable a number of hypotheses, derived from the theory, to be tested. These concern both the individual designer, the effect of the design process on his judgements and the reasoning behind his evaluations, and between designers, the levels of agreement between their priorities and between their evaluations.

The results of testing the hypotheses allow conclusions to be drawn about the validity of the theory. Generally the theory was borne out by the findings. Designers' verbally and numerically elicited values did provide logical and coherent explanations of their design proposals, and of their preferences in the evaluation of alternatives. Differences between designers' proposals and preferences could be explained by reference to differences between their elicited values. The judgement analysis technique chosen proved to be most useful, and checks within the experiments provide an assessment of its reliability in representing highly subjective judgements.

The thesis concludes by summarising the implications of these findings for research, for teaching and for practice.

CHAPTER 1 INTRODUCTION

1.1 Objectives

This thesis is intended to be a contribution to the theory of design. By theory is meant an exposition of the abstract principles which underlie the activity. By design is meant the process of contriving a man-made object. By theory of design therefore is meant an exposition of the abstract principles which underlie the process of contriving a man-made object.

The boundaries of design have been drawn in ever widening circles during the last three decades of concentrated interest in the nature of the design activity. There is a strongly held belief that the intellectual process underlying the design activity is common to a wide range of disciplines, including not only architects and engineers but also planners, managers, legislators and many others (Archer, 1969, 1974; Simon, 1969; Jones, 1970). Research into this intellectual process thus benefits from making use of the findings of a number of disciplines, and at the same time may contribute to their study.

In this thesis architectural design has been chosen as a focus for the study. The reasons for this are (after Jones, 1970; Mallen & Goumain, 1973, 1976):

1 Historically there is a long tradition of architectural theory recording amongst other things the influence of contemporary philosophy and science on architecture (Collins, 1961; Giedon, 1941; Wittkower, 1962; von Simpson, 1956).

2 Architecture provides shelter for people's activities and the backcloth to everyday life; where in the past buildings, streets and cities grew up anonymously, and architects provided a few landmarks of achievement, during this century there has been a wider spread of architectural forces in the environment. Complete new towns have been built, with decisions as wide ranging as from city plans to door knobs taken by designers.

3 Buildings are expensive and represent long term investments. They continue to shape people's lives long after they were first conceived.

4 Almost all buildings are permanent prototypes. There is rarely the opportunity to make changes in the light of experience; the building has to be 'right' first time.

5 The requirements for new buildings are constatly changing; higher standards of safety and comfort are expected, and statutory legislation increases to reinforce safety requirements and to reflect national policies.

6 Available materials change, technology changes and the range of potential solutions is widening.

7 New technologies and deeper understanding of existing ones result in a rapid growth of technical information to be assimilated by the designer.

8 Increasing scale and complexity of building proposals necessitate teamwork, with its inherent problems of management and communication.

9 The move towards public participation in design decisions implies the need not only for design intuition to be supported by objective and analytical means of design, but also for designers' assumptions to be stated explicity rather than occur implicitly.

10 Significant changes in the pattern of architectural teaching, in particular the replacement of articles of apprenticeship by university training, have necessitated making explicit the process of design in order to transmit it.

11 The reluctance of architects to use computer based design aids, by comparison with the enthusiasm of the engineering professions, appears to be attributable to the nature of the architectural design process and has already generated much interest in it.

Although not all of the above points are exclusive to architectural design, much of the recent literature on design methods and design research has concentrated on architecture. The body of knowledge being accumulated thus has the advantage of being based on a major design discipline and at the same time has implications for other areas of design activity. The research reported here represents an attempt to improve and to inform the understanding of the architectural design activity, and to contribute towards the debate about appropriate methods and techniques for developing this understanding.

In this thesis the sketch design process of individual designers is studied. The argument is proposed that design is a form of decision making, and that value theory may provide a theoretical foundation for the study of design. From the results of existing research a tentative theory is proposed. This is then tested in experiments, where groups of designers prepare sketch designs and, in conjunction with the design process, perform a judgement analysis exercise developed by Saaty (1977). The designs produced are evaluated by all the designers. From the data a number of hypotheses are tested about designer's values and the evaluation of designs.

The purpose of the research is to encourage the explicit formulation of decision making in order to promote individual designers' self-awareness of judgemental processes and to enable decisions to be explained, recorded and defended. Not only is it the belief in design research that awareness in action may lead to improvement, such a belief has been demonstrated in a teaching situation. Abercrombie (1960), working with a group of medical students, had them compare their judgements of, for example, X-ray plates. The varying descriptions made some of the factors influencing the judgements become apparent; for example, it encouraged the students to differentiate more explicitly between observations and inferences. From the results of a test at the end of the series of sessions, using her group and a control group, she was able to support the hypothesis that

> "we may learn to make better judgements if we can become aware of some of the factors that influence their formation." (Abercrombie, 1960)

This chapter goes on to describe the background to the thesis, outlining the author's experiences which stimulated the research. A previous dissertation (Lera, 1976) had generated a broad interest in design methods and theories. This interest led to a programme of research being started at the Department of Design Research, Royal College of Art. The first part of the research programme comprised familiarisation with the ideas culture of the Department. This, together with involvement in the monitoring of the practical use of a computer aided building design system by architectural design teams, gave rise to the tentative theory for exploration and testing. The exploration and testing of the theory forms the main body of the text. The next section of this chapter discusses these aspects in more detail, and the final section explains the structure of the thesis.

1.2 Background

The approach taken in this thesis has been influenced by three major factors. All three have directed the choice of research undertaken.

1.2.1 Author's previous dissertation

Between 1974 and 1976 the author undertook to write a dissertation about design theory, which was submitted in partial fulfilment of the requirements of the degree of Bachelor of Architecture at the University of Liverpool. Entitled The Architecture of Design, the dissertation began by considering the reasons for the dissatisfaction with traditional methods of design and the growth of interest in the design process. It went on to show that this growth of interest had led in particular to proposals for introducing systematic techniques and the principles of scientific method into design. A brief account was given of scientific method.

The main body of the work was a critical review of the design methods proposed in the 1960's. The inferences to emerge from the review were that many of the contributions were potentially fruitful in increasing knowledge of the design process, and improving design. However their presentations in the private languages and notation of different disciplines, and their often rigidly prescriptive nature, were as much a hindrance as a help. Reasons for their lack of acceptance by the architectural profession were explored, and felt to be due to a lack of compatibility with intuitive design processes.

The dissertation concluded in the firm belief that, for the benefit of design teaching and for the provision of design aids, it was essential to improve the understanding of the design process. Four studies of monitoring architectural design processes were chosen, and the dissertation ended by comparing them: Eastman (1970) and Foz (1972) had monitored designers by direct observation, Willey & Yeomans (1974) had analysed a sequence of sketch scheme drawings, and Krauss and Myer (1970) reported a case history.

The differing time scales, design problems and methods of monitoring made the comparisons difficult. A relatively superficial overview was given by combining the findings of the four studies, but to emphasise the need for further empirical studies the closing statement made the point that

> "The very diversity of the nature of the studies is a hindrance to their comparison and collation.... Only when experimental results can be meaningfully correlated can we attempt to comprehend the architecture of design."

1.2.2 Department of Design Research

After spending a year in practice with Messrs. Derek Phillips Associates and qualifying as an architect, the author began a research programme at the Department of Design Research in 1977.

Here the early work by Archer had been at the forefront of the design methods movement. His check-list type model, Systematic Method for Designers (1964)

> "turned out to be very helpful to quite a lot of designers and hardly a week goes by even today without my receiving a request for copies" (Archer, 1979).

Later, his The Structure of the Design Process (1969) set out a framework intended to form the basis of a science of design, and to be compatible with the disciplines of management science and operational research. Although apparently it was

> "never accepted by working designers in quite the same way" (Archer, 1979)

it did emphasise the importance of values in the design process, and introduced a formal numerical technique for evaluating alternative designs. Subsequently Archer (1971) provided a more detailed account of the nature of values.

More recently, in answering 'Whatever became of design methodology?', Archer has stated that

> "Design methodology is alive and well, and living under the name of Design research" (Archer, 1979).

Despite this assertion there seem to be significant differences between design method and design research. Method implies 'orderly procedure', and has the same root as methodical. Generally speaking design methods were orderly procedures prescribed for designers to follow. By contrast, design research is conducted, as the name suggests, in the spirit of inquiry.

It is the spirit of inquiry which characterises the more recent work in the Department. In particular a long term study commissioned by the Science Research Council in 1972 was undertaken by Mallen, Goumain & Purcell. This sought

> "to develop models of architectural design processes both as carried out by the individual designer and by the design organisation" (Mallen, Goumain & Purcell, 1974).

The need for such a study was due primarily to commitments to the development of computer aided building design systems, which demanded knowledge of, and sensitivity to, user requirements and methods of working. These demands

> "made it more than ever necessary to be explicit about such design processes" (Mallen et al, 1974).

The development of models of the design activity relied both on the case history approach using live building projects, and on observational studies in laboratory experiments. At first these studies were characterised by their concentration on observation and description; there was little underlying theoretical background from which to formulate hypotheses and to test them. Nevertheless the data obtained constituted accurate reliable records of design as practised.

Subsequent progress was due to the assimilation by the project team of current ideas and theories about human information processing in psychology, cybernetics and artificial intelligence. From these disciplines arose the notion that the design activity may be associated with the development of a model or internal representation which the designer has of the design problem and its content. Increasingly sophisticated techniques, derived from cognitive psychology, were used to elicit these internal representations. Kelly's (1955) repertory grid was used to elicit the subjective views of design problem structure, and multi-dimensional scaling technique used to infer structure from the repertory grid data (Mallen & Goumain, 1973; Stansall, 1973).

Two models of design activity were proposed. SIMDAC (SIMulation of Design ACtivity) purports to describe the structure of the individual's information processes as he designs. SHADO (Simulation of a Hypothetical Architectural Design Organisation) describes the operation of a design organisation in which specialists co-operate to carry out a number of different design projects.

As a model SIMDAC was not intended to be prescriptive in the tradition of design methods, but to be explanatory. In providing a simulation of designer behaviour

> "The resulting computer model will be directly testable. That is, it will produce sequences of behaviour which will be comparable with sequences of real life design behaviour" (Mallen & Goumain, 1973).

The model is described in detail by Mallen & Goumain (1973). Attached to this project, both Henrion (1974) and Cornforth (1976) monitored designers, using various techniques and methods of analysis, and thereby provided constructive criticisms of the SIMDAC model. These experiments are referred to in chapter 2.

It can be seen that the Department has been developing a consistent theme and set of principles in studying the design activity. The approach has been methodical in application, so that each new study can build on what has gone before. Furthermore, dedication to this subject matter has resulted in a network of contacts with other design researchers and periodic visits to appraise their findings (e.g. Mallen, 1978).

In the context of the present thesis it is significant that the Department's interest in descriptive models of the design activity had led Stansall, Henrion and Cornforth to consult many of the same references as the present author had in his dissertation, in particular studies which monitored design processes by Eastman and Foz.

1.2.3 Monitoring the use of a computer aided building design system

Concern for knowledge of, and sensitivity to, the design process, for the purposes of developing computer aided building design (CABD) systems, implies evaluating these systems to ensure that they reflect this concern. CEDAR3 (Computer aided Environmental Design, Analysis and Realisation) is a CABD system, which in its early phases had been developed in the Department of Design Research. Later, in 1973, it transferred to the Department of Environment Property Services Agency. CEDAR has been the subject of two formal evaluations. The first was of an earlier phase of the system (Thompson & Hughes, 1974). In 1978 the CEDAR3 system underwent pre-production trials intended to measure its efficiency, reliability and acceptability. The pre-production trials were held initially in collaboration with the Department of Design Research. Because of the unique opportunity which they presented for studying the (computer aided) design process the present author asked to become involved, and subsequently did so.

The trials took two forms. Short intensive design exercises (IDE's) were held, in which design teams, away from their normal place of work and with the usual office distractions and interruptions removed, undertook artificially constructed design projects and prepared sketch plans within, for instance, two days, with the computer system available to aid them. The system was also installed in DoE Project Offices for up to three weeks at a time, and building design teams with projects at the sketch plan stage were encouraged to use the computer to evaluate their proposals.

Both types of trial were monitored. Two approached were used. Direct monitoring of the process involved an observer (in most cases the author) following the proceedings by completing monitoring forms. On these were recorded the major design activities, cross referenced to duplicated copies of all computer input and output, and at the IDE's, to tape and video recordings. In addition to the direct monitoring the designers were asked to complete questionnaire surveys of their attitudes towards the use of computers in the context of building design, both before and after having experience of the system. And the designers were also asked to estimate the potential benefits of CEDAR3 in terms of their subjective probability estimates of savings in capital and running costs resulting from the facility to analyse rapidly alternative building configurations.

The trials were successful in providing data to enable assessments to be made of the efficiency, reliability and acceptability of the system. The evaluation was reported in an internal PSA document, a conference paper, and elsewhere (Thompson, Lera, Beeston & Coldwell, 1978, 1979; Thompson, 1979; Lera & Thompson, 1980).

It had been expected that most of the data would be taken from the duplicated computer input and output, with the monitoring forms providing detailed support where necessary. Such was the case for assessing how well the system worked in terms of software errors, operator errors, machine crashes and computing costs. However in attempting to show the expected benefits to building designs by studying the sequence of changing design variables (overall block form, window areas, U-values, etc.) it was found that these variables were only rarely amended systematically by the designer. From a few unmethodical analyses designers seemed to extrapolate implicitly. Where methodical procedures were used they were reported in the case studies mentioned above. But for the most part even the detailed monitoring carried out at the IDE's did not give a record of the full deliberations. It was necessary to rely on subjective probability estimates of benefits. This remains true not only of non-quantifiable benefits like improved daylighting of interiors, but even of thoroughly quantifiable benefits such as savings in capital and running costs.

In addition to the findings about the CEDAR3 system, involvement in the monitoring of design teams provided an excellent opportunity for gaining insight into some of the problems faced by research into the design activity. In particular the findings served to stress the need to elicit subjective data from the designers in a more precise way than passive observation can provide. They emphasised that some kind of judgment analysis technique was necessary to allow definite conclusions about benefits to be drawn.

1.3 Structure of the Thesis and Numbering System

This introduction has attempted to summarise the author's cumulative experiences of design method and design research which have led to the present study. It has sought to show that there are many sound reasons for studying architectural design.

Many of these reasons themselves contributed to the development of design methods and computer aids. However such aids have not been taken up by architects to any degree. This is believed to be due to lack of compatibility with the intuitive design process, and has stimulated research into the design activity. Such research is leading to developments of descriptive models of the design activity. The techniques being used to study the design activity include analysing sketch drawings, passive observation of the designer at work, asking the designer to 'think aloud' as he designs stream-of-consciousness style, through to interviewing him, asking him to complete questionnaires or to perform psychological measurement tests.

The thesis goes on to propose a contribution to the development of these models. Chapter 2 puts forward a tentative theory that the design activity may usefully be understood as one of exercising judgement, and that the resolution of inherent conflicts during the design process can be made only by recourse to value judgements, to deciding priorities among competing objectives.

Chapter 3 describes the programme of five experiments designed to test the tentative theory proposed.

Chapter 4 describes the experimental techniques used to elicit and to analyse designers' judgemental processes in the experimental programme.

Chapters 5, 6, 7, 8 and 9 each report an experiment: the hypotheses which the theory leads to, the methods used to test the hypotheses, the results found and the conclusions which may be drawn from the results, are given in each case.

Chapter 10 summarises the findings of the thesis. The tentative theory is discussed and the implications of the findings for research, teaching and practice are considered.

The numbering system in the thesis is intended to work as follows:

Subsections of the thesis are referred to in the text by underlining: this is subsection <u>1.3</u>.

The data from each experiment are shown in numbered tables and figures. These are numbered according to the chapter to which they refer, e.g. table 7.5 is the fifth table of chapter 7. The figures and tables appear at the end of the chapter to which they refer. Statistical analyses of the data appear within the body of the text; in some cases these analyses appear in tabular form.

Appendices are also numbered according to the chapter to which they refer, but they appear at the end of the thesis; thus appendix 6.1, for example, relates to chapter 6.

Those who took part in the experiments are referred to as subjects. The subjects are numbered according to the number of the experiment in which they took part, <u>not</u> according to the chapter in which the experiment is described. The subjects' numbers are prefixed by the letter S. The subjects in Experiment Three, for example, are S3.1, S3.2, S3.3, S3.4, S3.5 and S3.6, and this experiment is written up in chapter 7. The sketch plans produced in Experiments Two, Three, Four and Five are labelled by capital letters (in Experiment One the sketch plans had been drawn up previously). The alphabetical ordering of the letters corresponds with the numbering of the subjects; for example, S3.1 produced plan A, S3.2 produced plan B, and so on.

CHAPTER 2 DESIGNERS' VALUES AND THE EVALUATION OF DESIGNS

2.1 Introduction

A number of simplifying models to represent the complexity of the design process have been proposed during the last three decades of interest in the design activity. These models have been concerned with a range of issues pertinent to design; the organisation and management of a design project (RIBA, 1967), the morphology of design problems (Norris, 1963), the widening of the solution space through stimulation of the designer (Osborn, 1957), the setting of a series of stages for the designer to follow (Thornley, 1963), the problem solving processes of the designer (Eastman, 1970), the appraisal of designs (Markus, 1969), and the setting up of a framework for a science of design (Archer, 1969). Many models fall into several of these categories.

Several reviewers have provided ways of classifying these models. Jones (1970) presents a textbook classifying models according to their purpose and presenting them as practical problem solving tools. Lera (1976) divides them broadly into design-process based and design-product based models. Henrion (1974) separates descriptive from prescriptive models; normative models prescribe how design should be conducted, an approach typified by design methods; descriptive models purport to show how design is conducted, and are more typical of design research. A third approach is the chronological one; Broadbent (1977) describes three generations of design methods, each of which differs in its emphasis on who is to be responsible for design conjectures and who is to make the final decisions about the value of those conjectures. Gasparski (1979a) presents a tentative overview of the last three decades of design research, reviewing both theoretical work and empirical studies. Gregory (1979b) tabulates and classifies thirty observation-based studies of designing and asserts that they constitute the prime material upon which development of knowledge about designing can be founded.

The present thesis is a study of the sketch design process of individual designers. The argument put forward is that design may usefully be understood as a form of decision making. This chapter goes on to develop this argument. A typical design problem is introduced and analysed in terms of its dominant attributes and their relative values. Criticisms of the role of value judgements and of the conscious deliberation of relative values are themselves criticised. Existing descriptions of the design activity are cited, together with the analysis of the design problem exemplar, to demonstrate that value theory may provide a theoretical foundation for understanding design. The chapter concludes by presenting a tentative theory for exploration and testing.

2.2 A Typical Design Problem

Consider a typical example of the kind of design problem solved almost daily by most architects: the design of a domestic window [1]. The design problem may be represented by a set of attributes (Reitman, 1964). Attributes are those qualities which it is hoped will be attributed to the final design. Attributes in design problems are commonly stated in the form of imperatives (Simon, 1969).

The attributes might typically [2] be for the window that

it should provide a good view to the outside,
it should allow sufficient daylight in the room,
it should allow adequate ventilation,
it should have a pleasing visual appearance,
it should not result in excessive heat gain or heat loss,
it should not exceed a certain cost.

---

(1) For the purposes of illustration the presupposition has been made that the solution will be a window rather than, say, the provision of artificial light and mechanical ventilation, which would satisfy at least some of the attributes. The choice of example is supported however by the point that architects are frequently given design problems in this form, e.g. they are asked to design a school rather than to provide a means of education, a house rather than an organisation for coping with domestic activities. Of course they are still at liberty to question whether a window (school, etc.) is the best answer and to propose a solution which satisfies the attributes even though it might not be termed a window (school, etc.).

(2) A complete list of attributes for a design problem would be unusual. Certain attributes are commonly taken for granted. For example, the attribute 'the window must prevent the ingress of rain' is a quality which all windows ought to have. Therefore attributes in the sense used in the text refer to possibilities rather than absolute necessities.

In designing the window the designer specifies the decision variables, which might typically be

size,
proportion,
material of frame,
type of glazing.

Faced with this type of problem the designer knows that there are numerous alternatives each of which will result in different levels of fulfilment of the attributes. For example,

> a large window will give a good view, and good daylight, but it may be out of keeping with the room interior and the exterior; it may cause excessive heat gain and heat loss. Double glazing will reduce heat loss but will add to the capital cost. The large size may necessitate a costly material for the frame.

or

> a smaller window may be cheaper in capital cost, be more compatible visually both inside and out and not cause excessive heat loss and heat gain. But it may restrict the view, give little daylight in the room, and allow insufficiant ventilation.

In designing the window the architect, whether explicitly or not, is making a value judgement about the relative importance of the attributes. Through his choice he is ascribing different weights to the attributes according to the degree to which he values them. He may, for example, choose a window in which all the attributes are represented approximately equally, or he may value a splendid view and, taking that as the major attribute, ascribe only low weights to the others.

The design process inexorably entails the designer making decisions, either alone or through collaboration with his client and/or consultants, either explicitly or more likely implicitly, about the attributes he believes to be important and their relative weights. As Canter (1977b) has said

"The architect has to juggle the priorities".

The designer explores the problem and proposes solutions intended to satisfy the priorities he has decided. During the process of exploration unforeseen critical interdependencies may become apparent as the designer learns more about the problem. Thus during the evaluation phases of the design process the designer may or may not find that his proposals reflect his priorities. If they do he may move on to another part of the problem. If they do not he may feel that his proposal nevertheless represents a good solution and adjust in his mind the original weightings to correspond with what he has proposed. What the designer wants depends upon what he finds he can have. This process has been described by Frishmuth & Allen (1969) and has been termed 'solution by negotiation' in contrast to 'solution by innovation' by Archer (1969).

There is no such thing as a right answer to this kind of problem. There are many answers and they demonstrate each of the attributes in varying degrees. What is recognised as a good answer is dependent upon value judgements. In evaluating alternative proposals preference will be shown towards those which most closely reflect the priorities of the evaluator.

2.3 Critics of Value Judgements in Design

The importance of values and evaluation in the design process seems self-evident. Surprisingly there is a school of thought which has sought to deny this:

> "We believe that it is possible to define design in such a way that the rightness or wrongness of a building is clearly a question of fact, not a question of value" (Alexander & Poyner, 1970).

In developing a pattern language for design Alexander decided to regard all human tendencies as worthy of fulfilment and conflicts as the occurrence of tendencies coming into opposition as a result of inadequate forms. Patterns would allow tendencies to coexist without conflict. In A Progress Report on the Pattern Language Duffy & Torrey (1970) reaffirm the relationship of patterns to values

> "any approach based on the idea of the compromise of values or trade-offs is antithetical to the pattern language which attempts in each situation to achieve the best of all possible worlds by resolving all conflicts."

Daley detected serious philosophical inconsistencies in Alexander's beliefs and suggested that, although he claimed to be observing conflicts which were brought on by inadequate forms, more often

> "he seems to be defining conflict in terms of his own preconceived ideas about what constitutes bad form" (Daley, 1969).

She also noted that the question of observing tendencies in the environment could not be objective; tendencies might be undesirable, and conflict among tendencies might coexist within the same person. Thus Alexander

> "would have to decide which of the tendencies was worthy of fulfilment or facilitation in by the environment, and that sort of decision, which would surely crop up repeatedly in any realistic assessment of human conditions, requires an appeal to values beyond Alexander's simplistic fiat that the sole criterion of 'rightness' in environment is the fulfilling of human tendencies" (Daley, 1969).

More recently March has made a detailed examination of some of the unwritten assumptions and inaccuracies in Alexander's derivation of patterns.

First March shows that statements about conflicting tendencies

> "are about values. Each can be rewritten 'X prefers ...' and is therefore a statement about preferences. It is always possible to give such preferences a partial ordering and the design task can then no longer avoid the problem of evaluation" (March, 1976b).

Second he shows that whereas Alexander puts forward one solution (pattern), justified by an ostensibly scientific explanation, on a take it or leave it basis, that this is an example of 'false precision' and that a more rational attitude leads to the selection of

> "a solution from a range of possibilities and attempts to assess its relative value" (March, 1976b).

Through his examination of the logic of design and the question of value March is unequivocal that

> "value theory is the essential foundation of any rational theory of design" (March, 1976b).

2.4 Critics of Weighting in Design

That attributes are weighted differently may seem to be self-evident, but there are those who have criticised ranking and weighting procedures. Jones (1970) characterises attempts at weighting as absurd, and according to Grant (1974), Alexander & Manheim (1962) also argue that consciously deliberated weights are not valid.

In support of his case Jones notes that in order for numerical weights to be assigned the data must be measurable on an interval or ratio scale. He also describes the problem of intransitive relationships in the process of ranking and the requirement that the attributes have to be independent for weighting procedures to be theoretically valid.

Grant however makes the point that

> "people must and do make decisions in multi-criterion situations, and act on them, whether or not the decision situation is theoretically well behaved and whether or not various criteria can be demonstrated to be factually independent" (Grant, 1974).

Grant has described three options which those who do not believe in the validity of weighting procedures may follow in order to make essential judgements.

1 Find the one most important attribute and decide on the basis of that attribute alone. This is of course a form of weighting albeit rather simplistic. The main implication of this method is that it involves hoping that all other attributes are satisfied at least to an acceptable degree. In any event, as he notes, this may be regarded as a form of weighting in which all attributes bar the main one are zero rated. Grant concludes that this approach is an inadequate response to a complex problem.

2 Attempt to construct compelling graphic layouts from the partial judgements. Grant (1974) describes two attempts to employ map overlay techniques of decision criteria for highway locations. Each shaded overlay represents a decision criterion, and thus when all are overlaid the resulting shading indicates the optimum route. The technique has apparently been used by Alexander & Manheim (1962) and by McHarg (1969). According to Grant, McHarg simply combined the maps and thus built in an implicit equal weighting to each criterion. Grant asserts that Alexander & Manheim however did not merely overlay all the maps representing the decision criteria,

> "they combined similarly patterned maps into one representative composite for each set of similar patterns. By so doing they assigned accidental weights of importance to each map or decision criterion and the accidentally assigned weights varied widely in magnitude" (Grant, 1974).

Grant goes on to quote an example in which he claims one map was weighted 62.5 times as heavily as another by graphic accident and without intelligent deliberation.

3 Consider all criteria carefully, then sit back and let the matter incubate and an implicit intuitively derived decision may emerge. Grant asserts that even here

> "It can be argued that the process of deliberating

> and aggregating weighted partial judgements is in fact a model of the process that one's mind must go through in arriving at an intuitive implicit gestalt decision, as a result of considering multiple criteria and then allowing an ensuing period of incubation" (Grant, 1974).

Just as Daley and March showed that critics of the notion of values and value judgements merely made the value judgements implicitly and then disguised them under the claim of factual objectivity, so Grant has shown that critics of weighting procedures in proposing alternatives have been known to make weighting decisions by default or implicitly without explicit recognition or deliberation.

2.5 Designers' Judgement

> "Le raisonnement, la critique, viendront à leur tour pour contrôler votre conception, car après avoir imaginé il faut que vous sachiez être les propres juges de votre imagination" (Guadet, quoted by Collins, 1971) (1).

> "In the design process judgement constitutes one of the integral creative components, in that it is the mechanism by which the relationship between intuitively imagined forms and intellectually apprehended data is continually assessed. For reasons stated earlier this aspect of judgement can be most conveniently considered in terms of 'decision making' because although, in theory, it would be possible for an architect to complete several different projects for any one building, and then 'judge' which is the best, in practice the process of selection can usually be effected most efficiently at enbryonic stages in the course of the design, whereby only one final project is produced" (Collins, 1971).

Judgement, 'deciding the merits of', and evaluation, 'determining the value of', have, with the exceptions noted above in 2.3, been widely accepted in prescriptive and descriptive models of design. For example, a number of design methods were based on the three phase cyclical process: analysis - synthesis - evaluation.

---

(1) "Reasoning and criticism come in turn to control your ideas, because having used your imagination it is necessary to know how to exercise proper judgement of it" (author's translation).

> "One of the simplest and most common observations about designing and one upon which many writers agree, is that it includes the three essential stages of analysis, synthesis and evaluation.... Most design theorists agree that it is usual to cycle many times through this sequence" (Jones, 1970).

More recently Hillier, Musgrove & O'Sullivan (1972) have proposed a new paradigm for design. They argue that

> "design problems are essentially pre-structured both by constraints and by the designer's own cognitive map.... Design proceeds by conjecture-analysis rather than by analysis-synthesis."

Later they write of conjectures

> "By and large they come from the pre-existing cognitive capability - knowledge of the instrumental sets, solution types and informal codes, and occasionally from right outside - an analogy perhaps, or a metaphor, or simply what is called inspiration."

And of analysis they write

> "the purpose of analysis is primarily to test conjectures."

It is suggested that analysis is perhaps not the best term in this instance; 'testing conjectures' implies 'deciding the merits of' or 'determining the value of'. If this suggestion is accepted then the account by Hillier et al would be in close agreement with those quoted from Guadet and Collins at the beginning of this section. Design may be resolved broadly into imaginative and evaluative forces. Again therefore evaluation plays an essential role in design.

In descriptive models derived from observations of designers there is both general recognition and detailed description of the role played by evaluation.

Lawson's (1972) results are strongly supportive of the conjecture-evaluation paradigm in design. He studied strategies used in two dimensional spatial layout problem solving by architectural students and science based (non-architectural) students. In comparing their strategies he found that whereas the science based students tended to search for underlying rules (analysis) and then propose a solution which satisfied those rules (synthesis), the designers proceeded by trying alternative configurations (conjecture) and testing whether they complied with the rules (evaluation). He described the former strategy as problem-focussed, the latter as solution-focussed.

From his monitoring of designers planning a bathroom layout Eastman (1970) drew a similar conclusion about generative and evaluative forces in design.

> "Instead of generating abstract relationships and attributes, then deriving the appropriate object to be considered, the S's (subjects) always generated a design element and then determined its qualities."

Foz (1972) monitored four subjects of varying degrees of design training during a two-hour architectural sketch design problem. His findings support the same contention. He argues that the design activity proceeds as ad hoc responses to perceived misfits between a 'pre-solution model' evoked from memory and the program (design) requirements. Both Eastman and Foz applied an information processing theory of cognition to help to provide explanations of their observation-based studies in terms of cognitive processes in design.

As Gasparski (1979a) has noted, in addition to observation-based studies of designers a new and promising trend in design research is the effort to identify the internal representations used by designers. This approach has been proposed in the study of the design process at the Department of Design Research, as noted in chapter 1. Mallen & Goumain (1973) citing psychologists Piaget and Bruner and other research in Artificial Intelligence and Heuristic Programming, posit the hypothesis that

> "just as the child develops and uses internal representations of increasing sophistication to gain control over his environment, and as the master chess player uses a powerful representation to avoid exhaustive search in chess, then so does the designer develop and use internal representations of design problems to organise and control his progress through the design task."

They argue further that the internal representation is a dynamic plan of action for dealing with the problem. The SIMDAC model they propose (see 1.2.2) is intended

> "to simulate the operation of internal representation processes" (Mallen & Goumain, 1973).

The fundamental mechanism of the model is the cybernetic feedback loop described by Miller, Galanter & Pribram (1960) as a Test - Operate - Test - Exit, or TOTE, unit.

In collaboration with the study by Mallen & Goumain three research workers explored further the implications of the hypotheses proposed. Stansall (1973) used Kelly's repertory grid technique to elicit the form of designers' internal representations. Henrion (1974) observed designers in order to describe the nature of their plan of action in solving a design problem. Cornforth (1976) combined these two approaches, using repertory grid technique and multi-dimensional scaling (MDS) analysis of the data to elicit designers' internal representations, and observing designers solving a sketch design problem. He then attempted to compare the internal representation revealed by the MDS analysis with the plan of action, as observed in the designer's strategy.

Stansall in using cluster analysis of the repertory grid data found that experienced architectural students revealed a greater number of separate clusters of constructs than did inexperienced architectural students.

Henrion (1974) monitored four subjects, two designers and two non-designers, arranging furniture in an office layout. His study of verbal protocols obtained from the designers dealt primarily with the way constraints operated. He studied how conflicting constraints were identified, before or during the process, and how they were resolved, partially resolved through compromise or not resolved but accepted. He presented some of his findings in the form of a graph of constraints identified and satisfied

> "intended to be a simple model of the subject's changing evaluation of the arrangements he generates in terms of the number of constraints it satisfies."

He characterised the design process as a series of modifications to the initial layout during which successive layouts satisfied increasing numbers of constraints. He concludes by stating that

> "the design process was better modelled as a continuing attempt to increase the number of satisfied constraints, although it is clear that no solution exists which can satisfy them all" (Henrion, 1974).

Cornforth (1976) set up an experiment in which designers were monitored 'thinking aloud' while undertaking a sketch design scheme, and in conjunction with the design process completed a repertory grid. Thus a verbal protocol could be transcribed and the results compared with a multi-dimensional scaling analysis of the repertory grid data. He characterised the sketch design activity as a combination of a specification process and a search process. He offered evidence to support the hypothesis that search takes place in a hierarchy of problem spaces, and that the trend in the design process is to work from general simplified representations to more specific detailed representations. This is clearly apparent from the verbal protocol which Cornforth provides in the appendix. Right at the beginning of the design process the designer takes an overview of the problem and proposes an outline solution, which is then successively modified as new constraints are identified. He noted that constraints were identified in two ways: from memory and through perception of a misfit in a configuration. However the comparison between the plan of action and the internal representation proved to be of limited value:

> "No structure could be found in the (MDS) configurations which corresponded to the detailed behaviour of the subjects."

Nevertheless considerable insight was gained into the design activity, and a number of observations made about the SIMDAC model. The experimental findings were in general agreement about SIMDAC, though a number of modifications were suggested. Cornforth, like Henrion, was unable to detect the relative importance of the constraints he identified in the protocols.

Elsewhere Baer (1976) and Akin (1978a, 1978b) have also conducted research into the design process within the framework of an information processing theory of cognition. Akin's study of the architectural design process was conducted in order to propose a descriptive model of the design behaviour of architects. He provided evidence from protocol analyses of designer behaviour to support the existence of eleven different information processing mechanisms in design, and explored three of them, 'design plans', 'transformation rules' and 'design symbols' in some detail. Among his many conclusions about design strategies and information processing mechanisms, are several findings about designers' judgemental processes. On conflict resolution in design he notes

> "conflicts are resolved either by remodifying the physical description or by modifying the problem criteria" (Akin, 1978a).

He also provides evidence to support the conjecture-evaluation paradigm

> "Often a few cues in the environment are sufficient to evoke a pre-compiled solution in the mind of the designer" (Akin, 1978b).

Another important study of the designer's internal representation or conceptualisation of a design problem is reported by Aish (1974). He used connectivity analysis in the design and evaluation of a control console layout. He took one attribute only, adjacency or interaction of elements, and compared, using connectivity analysis, the degree of complexity of interaction of elements specified by the client, achieved by a clustering algorithm, achieved by the designer's conceptualisation (as elicited in a word association test), and achieved in the designer's proposed console layout. One of the more important findings was that the designer's conceptualisation achieved measurably less richness of interaction among elements than specified by the client, and that the designer's solution achieved measurably less richness of interaction than the designer's conceptualisation.

In addition to the use of psychological measurement techniques used to elicit designers conceptualisations of problems, and the observation of designer behaviour to study information processing mechanisms in the design process, researchers have shown the benefits of interviewing designers about their own design processes, or of listening to and interpreting their accounts of their own design processes. Although such techniques imply subjective interpretations of the data by the researcher, the two following accounts both indicate the relative importance of constraints, an aspect of the design process which few of the previous accounts cited had been able to express.

Darke interviewed a number of architects about their design process. She was able to provide strong support for the conjecture-analysis (or conjecture-evaluation) model of design. Furthermore she found a clear indication of architects' priorities from the interviews.

> "It has been suggested in this paper that designers do <u>not</u> start with a full and explicit list of factors to be considered, with performance limits predetermined where possible. Rather they have to find a way of reducing the variety of potential solutions to the as yet imperfectly understood problem, to a class of solutions that is cognitively manageable. To do this they fix on a particular objective or small group of objectives, usually strongly valued and self-imposed,

for reasons that rest on their subjective judgement rather than being reached by a process of logic. These major aims, called here primary generators, then give rise to a proposed solution or conjecture, which makes it possible to clarify the detailed requirements as the conjecture is tested to see how far they can be met" (Darke, 1979).

Where Darke reports on the designer's major aims as a small set of objectives, Grant suggests that the designer extablishes priorities among his objectives in a way analogous to weighting and ranking procedures. He reports listening to a talk by an architect in which

"he described in his own approach a process in which the various opportunities and constraints of the site and of the client's needs and desires were weighed and ranked just as effectively as is done in the systematic procedures familiar now. His personal design process was one in which carefully thought out personal decisions were effectively integrated into overall judgements that led to a most worthwhile house" (Grant, 1974).

There is one other approach to understanding the design process which, although it does not seem to have been made the subject of research in architectural design, has been used with interesting results to study computer programmers. Weinberg (1972) ran controlled experiments with computer programmers to find out how the specifying of different objectives or attributes would influence both the process and the product. Four programmers were given identical programming problems to solve, but two were asked for the program in as short a time as possible, the other pair that it should be as efficient in machine time as possible. The experiment was repeated with four other programmers. He found striking differences in the resulting programs, directly attributable to the different conceptions of the objectives; objectives not stressed were sacrificed to those stated explicitly. He found that the design processes varied too; different objectives caused different strategies to be followed by the programmers, particularly in their reaction to unanticipated difficulties. One of Weinberg's findings about computer programmers from his experiments is especially important in the context of the present thesis, for one of his conclusions was to suggest that a large proportion of the variation between programmers on any job can be attributed to a different conception of what is to be done; that is, programmers' differing values account in large part for the variations in their achievements.

To draw to a close this summary of accounts of designers' judgemental processes, mention should be made of some experiments which have been performed in the evaluation of designs. A large number of studies have been conducted under the heading 'architectural psychology' which attempt to establish user attitudes to buildings, and also to correlate the measurable performance of buildings with users' verbal responses. A number of techniques have been used in this research including Osgood's semantic differential (Canter, 1969a; Canter & Wools, 1970; Wools, 1969) and Kelly's repertory grid (Honikman, 1973). This research generally has not been conducted to study the design process, although Abel's (1975) 'Architrainer' was an attempt to teach students of architecture about their client's constructs using Kelly's repertory grid. However there are three studies in the evaluation of designs which are of direct relevance to the present study.

Lowe (1970, 1972) obtained evaluations of seven redrawn student architectural design drawings. The evaluators were lecturers in schools of architecture. The designs were evaluated with respect to two criteria, 'functional planning' and 'effective use of daylight and sunlight'. Evaluations were made individually with respect to the first criterion, then after a discussion between a group of three assessors (to simulate a school of architecture jury) individually with respect to the second criterion. The method of ordinal paired comparisons was used for the evaluations so that inconsistencies could be measured. Lowe found that most assessors were able to maintain a consistent criterion of preference through the assessment session. In assessing the levels of agreement between the judgements he found that there was significant concordance between the judgements with respect to each of the two criteria.

Cakin (1976) presented groups of people with five alternative design solutions for holiday chalets and asked them to put the designs in rank order of merit. The experimental subjects belonged to two categories: students of architecture and non-architects. The presentation of the schemes took two forms: crude information (plans and elevations) and sophisticated information (plans, elevations and performance profiles). Cakin measured the concordance reached by the groups. He found that

1 Groups of non-architects, given either crude or sophisticated information, exhibited significant concordance.

2 Groups of architectural students, given sophisticated information, exhibited significant concordance.

3 Groups of architectural students, given crude information, did not exhibit significant concordance.

Among his conclusions he suggests that

> "One explanation for the differences found between the agreement levels and preferences of judges could be that each individual has a mental profile consisting of cost, performance and other attributes, each attribute having a different degree of importance. If the attributes he is presented with are the ones he thinks important then his judgement will largely be influenced by the profile rather than by the drawings" (Cakin, 1976).

Later he writes

> "non-architects gave more consideration to the cost and performance measures of the schemes than the architects did. Architects seem to base their judgements on design drawings only" (Cakin, 1976).

These comments would certainly help to explain his results; the non-architects comparing primarily the given performance profiles would have an objective basis, the profile shapes, on which to make comparisons; the architects, not relying on the profile shapes and differing in the attributes to which they attach value, would therefore differ in their preferences.

Huber, Sahney & Ford (1969) obtained evaluations of twelve hypothetical hospital wards from thirteen senior hospital staff members. The wards were described in terms of seven quantitative factors. Each subject marked each of the wards on a 1 - 100 scale. Huber et al tried to fit the data to three forms of utility model and found that

> "arguments supporting the use of addilog or multiplicative model forms were not particularly relevant in this experimental situation."

They concluded that a subjective evaluation model of the linear type was as useful as the other two to represent judgements in such a case. A model of this type has been used in this thesis to represent judgements, as will be described in section 4.4. It is referred to in this thesis as an additive utility model; this is the generic term for such utility models.

The evidence so far cited in this chapter shows that even in the design of a modest window a decision implies a value judgement about the relative importance of attributes. Although some authors have criticised the conscious deliberation of value judgements, these views are shown to be invalid. Several authors are cited who emphasise the role of judgement and of evaluation in the design process, and a number of studies of designers provide empirical support for this contention; indeed Hillier, Musgrove & O'Sullivan create a kind of paradigm for design around the twin forces of conjecture and evaluation. Other authors have stressed that decisions are made with respect to a simplified representation, or have shown that the internal representation does not cope with the full complexity of the problem. One author has argued that designers fix on a small group of strongly valued objectives to generate their conjectures, and another that the designer establishes priorities among his objectives analogous to weighting and ranking procedures. This large body of evidence is strongly indicative of the need to refer to values in explanations of the design process. If value judgements play an essential role in the design process, then value theory may provide a basis for understanding decision making in design.

2.6 Value Theory

The link between values and their expression in decision making is well argued in value theory. Rescher, for example, in his Introduction to Value Theory describes values as being manifested through decision making, in words and in deeds, and he notes the difficulty of defining value other than by reference to these manifestations. But by observing actions and words, values may be inferred. Having a value is different from having a goal but the two are linked in that one's goals are reflections of one's values, he argues; the fundamental role of a person's values is to determine the evaluation of his actions and thereby to support practical reasoning. Practical reasoning encompasses rational deliberation in the assessment of alternative courses of action; the comparative assessment of alternatives in the search for the optimal choice among competing mutually incompatible courses of action can be made only by recourse to value judgements. He argues further that in the logic of practical reasoning, values are an essential component and provide criteria for choosing among courses of action that are mutually exclusive in the context of finite resources.

Tribus (1969) and Ozbekhan (1972) have given similar accounts to that by Rescher explaining actions, decisions and outcomes, and their relationship to value judgements. For example, Ozbekhan writes

> "(i) In order for man to act (rationally), a near or distant outcome must be visualised; (ii) such an outcome must be desirable; (iii) the desirability of an outcome can be judged in terms of its value, and the action leading to this outcome justified in terms of such value; (iv) if the actor has to choose among several outcomes, his preference for one particular outcome must also be justified with respect to its value; (v) choice among outcomes enters into the action equation only when there are alternative valued outcomes available; (vi) the spectrum of alternative valued outcomes corresponds to the spectrum of options available" (Ozbekhan, 1972).

The organisation of a person's values constitutes a value system (Bross, 1953; Rescher, 1969; Rokeach, 1973). Rokeach (1973) describes the function of a value system as a general plan employed to resolve conflict and to make decisions. He writes

> "Since a given situation will typically activate several values within a person's value system rather than just a single one it is unlikely that he will be able to behave in a manner that is equally compatible with all of them.... A value system is a learned organisation of principles and rules to help one choose between alternatives, resolve conflicts and make decisions."

This notion of a 'general plan' employed to make decisions is reminiscent of the 'internal representation' posited by Mallen & Goumain, as described above in <u>2.5</u>. Furthermore both seem to equate with the views of March & Simon (1958) about decision making.

> "Choice is always exercised with respect to a limited, approximate, simplified 'model' of the real situation."

Value theory, in addition to being a descriptive endeavour in philosophy and social science (Rokeach, 1973; Laszlo & Wilbur, 1973; Vickers, 1968), has also been developed as a formal numerically based theory of decision making. The seminal work in this area was <u>The Theory of Games and Economic Behaviour</u> by Von Neumann & Morgenstern (1947). In it they set out the conditions for a theory of value. Subsequently Luce & Raiffa (1957) gave a more general account of value theory and reformulated the set of axioms of rational behaviour.

From the theoretical issues which have been explored, a number of techniques for decision making have been gathered together under the headings of decision theory and decision analysis. Behavioural decision theory is the study of the way decisions are made; Edwards (1967b and 1967c) reviews the extensive research that has been conducted under this heading. The methods of decision analysis are derived from behavioural decision theory, but are intended to prescribe systematic frameworks within which decisions may be made. Keeney & Raiffa (1976) give an extensive account of formal techniques for making decisions with multiple objectives. Kaufman & Thomas (1977) provide a collection of papers illustrating applications of these procedures in planning and management decision making.

The formal study of decision making using techniques and theories developed in decision theory has not found application in architectural design. An exception is the work of Derbyshire (1976) who reports a study of indifference curves to represent the trade-offs made by architects and consultants between capital costs and running costs. However the view of design established in this chapter indicates that more than two attributes may be taken into account, and furthermore that these attributes will be of both a qualitative and a quantitative nature. An alternative technique, multi-attribute utility analysis, seems more apposite to the view of architectural design established here. Grant (1976, 1978) and Wise (1978) have recently discussed the theory and potential of multi-attribute utility analysis in design, though neither presents empirical evidence or examples of its having been used to explore designers' judgement.

Multi-attribute utility analysis entails the following points:

1 There is a set of alternative outcomes.

2 There is a set of attributes.

3 The outcomes demonstrate different degrees of fulfilment of the attributes.

4 The decision maker has a preference ordering among the attributes; he can assess the relative weights attached to the attributes.

5 The decision maker can assess the probability that any given alternative will fulfil an attribute.

6 The decision maker selects the alternative which maximises his utility function, that is, which in his subjective judgement fulfils those attributes which he most values.

It may be seen that this description corresponds to some accounts of design given in 2.5 and in particular it corresponds closely to the example of a typical design problem given in 2.2:

1 The set of alternatives are represented by the possible window designs.

2 The various window designs result in different costs, lighting levels in the room, heat losses, and so on.

3 The designer has a preference ordering among these attributes; he may value the view above all else or he may consider each of the attributes mentioned to be of broadly equal importance, for example.

4 In designing the window he bases his choice on achieving or fulfilling those attributes in proportion to the degree to which he values them.

Thus multi-attribute utility analysis may explain design decision making and may provide a suitable approach for studying design. According to this approach the designer may be considered to decompose the problem into the design variables and the attributes manifested by these variables. He assesses the subjective value or utilities of the attributes. He also assesses his expectation of the degree to which the choice of an alternative will fulfil a certain attribute. A folding back operation using the utilities and subjective probabilities of outcomes gives the subjective expected utility of each outcome. This subjective expective utility is the summation of the probabilities of alternative outcomes combined with the values attached to those outcomes. The designer's choice maximises his expected utility. The technique of 'prioritization' developed by Saaty (1977, 1978) provides both a means of eliciting this subjective structuring of problems in design, and a means of evaluating alternative designs, as described in chapter 4.

2.7 A Tentative Theory

This chapter has sought to show that judgement is an essential component in the design process. In assessing the reasoning behind judgement in the design process it has been shown that such judgements may be accounted for by recourse to values. Value theory provides a basis for understanding decision making during the design process. Multi-attribute utility analysis provides a framework for studying design.

Using the framework of multi-attribute utility analysis together with descriptive accounts of design, a tentative theory may be proposed:

1. Designers may initially rely on a small set of strongly valued attributes to generate their design conjectures.
2. The attributes designers value may be understood as being weighed and ranked; design decisions imply such preference orderings.
3. The weighting and ranking of attributes may change as the designer finds he needs to negotiate a solution. The internal consistency with which attributes are weighted may improve as a result of the design process.
4. Designers may differ in the attributes they value, and in their evaluations of the same attributes.
5. The differences between designers' value systems will account for the differences between their design proposals.
6. The differences between designers' value systems will also account for their differing evaluations of alternative design proposals; they will favour plans which reflect their own priorities and reject plans which do not.
7. Therefore if designers differ in their rating of attributes they will differ in their evaluation of alternative designs; if they are in agreement in the rating of attributes they will be in agreement in their evaluations of alternative designs.

From this tentative theory a number of hypotheses may be derived. An experimental programme, consisting of the organisation of five design and evaluation exercises was devised, and experimental techniques were selected, to enable the hypotheses to be tested. Subsequent chapters describe this programme and the results achieved.

It is of vital importance that the approach of stating and testing this theory does not rely on the inferring of values from the observation of design decisions. If it were to do so, much of the theory would be merely a presupposition. Although it is possible to infer values from decisions, the procedures demand highly constrained choice experiments and/or a large number of observations to ascertain the attributes used and their values. Observation of the design process does not provide suitable data.

The approach taken in the testing of the theory involves using a technique to elicit and to analyse designers' judgement. This approach lies between highly constrained choice experiments where, from the pattern of choices, values may be inferred, and open ended interviews which necessitate subjective analysis by the experimenter (Darke, 1979).

The purpose of testing the theory is to provide analyses of designers' judgement so that their behaviour, as manifested in their designs and in their evaluation of alternatives, can be explained by reference to their value systems in an objective and reproducible way. Therefore the thesis is intended to show not only that designs and designers' values differ but also to show how and why they differ. In this sense the thesis is additionally concerned with assessing the worth of the judgement analysis technique used for its ability to extract underlying structure from subjective data and thereby to provide meaningful, consistent results which permit explanations of designers' values and their evaluation of designs.

CHAPTER 3 AN EXPERIMENTAL PROGRAMME

3.1 Introduction

In order to test the tentative theory proposed in chapter 2, an experimental programme was devised. The theory gave rise to a number of specific hypotheses, and the experiments were devised in order to provide data to enable these hypotheses to be tested. The experimental programme comprised a sequence of five exercises in design and evaluation, and each exercise formed the basis of an experiment. The first, and to a lesser extent the second experiment, may be considered pilot studies, providing experience of the techniques intended to elicit the data, enabling an assessment to be made of these techniques, and offering some data for analysis. Confidence in the suitability of these techniques to furnish useful data led to the last three experiments. These were organised almost identically, their main difference being the choice of participating subjects: non-architects, students of architecture and qualified architects, respectively.

The experiments investigate designers' priorities and the evaluation of their design solutions. Experiment One was concerned only with the evaluation of alternative existing designs. Each of the other four experiments was based around an intensive design exercise (IDE). For these exercises, sketch design problems were set, intended to have a certain degree of realism, but at the same time not to be overcomplex for the time allocated. In conjunction with the design process the attributes which the subjects considered to be their priorities were elicited and scaled using 'prioritization' (see chapter 4). After having designed their proposals the subjects evaluated them, both with respect to the attributes elicited, and for overall merit. Precise descriptions of the running of the experiments accompany each one.

The organisation of the experimental programme, the number of experiments, the number of subjects and the analysis of results, were considered to be of an order compatible with the resources of time and money available to the experimenter.

3.2 Intensive Design Exercise

IDE's, as the name implies, involve compressing the decision making process into a much reduced period of time than would normally be allocated. The designers are sometimes removed from their normal place of working and the usual office distractions such as telephones are not allowed to interrupt. IDE's have been used extensively in recent design research as described in chapter 2. Although no comparative studies seem to have been carried out to establish the effect of the reduced time scale on design decision making, there are numerous advantages where monitoring is concerned. In particular, the lack of distractions and the short time scale prevent attention being divided and concentration lost. This is important when the designer is questioned about his priorities before and after the process; in an IDE the answers are spontaneous, there is less chance of rationalisation after the event and less chance of their being distorted by irrelevant occurrences. For example, if the designer's priorities change during the process, this must be recorded immediately at the end, for they may revert over time as the lessons of that particular problem are forgotten.

3.3 Experimental Conditions

Experimental conditions were held as constant as was considered compatible with the intended purposes of the experiments, and with the practicalities of a group of subjects taking part for a whole day or more.

In Experiment One each subject was interviewed individually. The interviews took place under informal conditions at the experimenter's work-space, the experimenter's home, and in two cases at the subjects' homes. As the experiment was conceived as a pilot study, intended to explore the potential of the techniques, it was not felt necessary to create, at this stage, sophisticated and uniform experimental conditions.

Experiments Two, Three and Four employed much greater control over experimental conditions. Each of the experiments was conducted in the subjects' studio: the students' normal design school studios in Experiments Two and Four, and a research studio in Experiment Three.

In Experiment Five each subject performed the experiment individually. In five cases this was done in two stages, both at the subjects' own homes. One subject, S5.2, performed the experiment, in two stages, at his office.

It can be seem that experimental conditions have varied within some of the experiments, and between experiments. Potential conditions for this type of experiment vary considerably. At one extreme Lowe (1972), in obtaining evaluations of architectural drawings, provided identical viewing conditions with uniform lighting levels, for his subjects to view the drawings. At the other extreme experimental conditions in obtaining designs from architects (Edwards, 1974) or in interviewing architects (Darke, 1979) have apparently been considered of such little importance that they are not even reported.

In the present experiments, because the subjects are expected to design as well as to evaluate proposals, the question of imposing standard conditions seems to necessitate a balance. It is at least arguable that to impose standard conditions such as drawing board type, drawing instruments, paper type and size, illumination levels and so on, might have a deleterious effect on the design process. This is especially true when the attempt is being made to elicit the subjective values of the participants. As an example, consider the observation of a fine artist at work; one would hardly propose to uproot him, together with his palette and easel, to a laboratory without expecting to affect his style. Therefore although the time-scales make the design exercises intensive, the experiments generally have been conducted in the subjects' normal working locations or design studios where they use their own materials and equipment. Only the non-architects in Experiment Three had to be provided with architects' scale rules and shown how to use them. In this way the experimental conditions may be said to lie between the very high degrees of experimental control employed by those such as Lowe, and the much less uniform conditions common in interview techniques. Differences between the experiments, both in the experimental conditions, and more particularly in their differing organisations, does warrant caution in the comparisons of results between experiments.

3.4 Number and Type of Subjects

The number of subjects whom it is possible to include is the result of a balance between attempting to penetrate the design process deeply, and ensuring that the subjects are not so few in number that their idiosyncracies obscure statistically generalisable results. Such a dilemma is commonly felt in design research. In the present study this balance has been made with primary regard to the choice of experimental technique and to practicality.

Experiments One, Three, Four and Five all use six subjects. The number was chosen for the following reasons:

a It gives a manageable group for discussion and agreement of common attributes, where those are used.

b When the subjects evaluate the designs produced by means of paired comparisons the number of comparisons, ${}^{6}C_{2}$, gives a reasonable number, 15.

c When the subjects perform similarity judgements of triadic comparisons the number of comparisons, ${}^{6}C_{3}$, also gives a reasonable number, 20.

Experiment Two was performed with the fourth year students at Liverpool School of Architecture and was tailored around the number of students; it differs in this respect from the other four.

The subjects themselves were chosen to represent three levels of architectural training. The three categories were architects, students of architecture, and non-architects. Experiment One included two subjects from each category, to obtain an indication of the effect that lack of architectural training might have on the answers which could be obtained. Experiment Two involved students of architecture. Experiment Three involved six non-architects. Experiment Four involved six students of architecture. Experiment Five involved six architects.

Experiments Three, Four and Five enable some comparisons to be made not only within the homogeneous groups but also between groups, although experimental conditions and organisation were not absolutely identical.

3.5 Time-table

The time-tables for Experiments Three, Four and Five were established by reference to the length of time taken to evaluate six design schemes in Experiment One. For convenience it was decided to try to concentrate the experiment into little more than a day for Experiments Three and Four, and two half day sessions with each subject for Experiment Five. This meant that the design process was allocated a period of about three hours, and this in turn was decisive in the choice of design task set. In Experiment Two the schedule for the exercise had already been established as a one week sketch design.

3.6 Design Task

The design task for all experiments except Experiment Two involved school planning. There were several reasons for this choice:

a Although schools vary greatly in their complexity it was felt possible to compile a brief for a school which would be realistic, would be of sufficient complexity not to be trivial, and would not be excessively difficult in the time allocated.

b All subjects will have a broad familiarity with the functioning of a school through personal experience.

c There is a precedent in other studies of the design activity at the Department of Design Research by Mallen & Goumain (1973), by Stansall (1973), and further unpublished work by Mallen.

d Design research elsewhere has concentrated on school design (Krauss & Myer, 1970; Willey & Yeomans, 1974).

School planning was therefore chosen. In particular the unpublished research undertaken by Mallen had used six primary school plans and these were reused in Experiment One. The success of Experiment One led to the adoption of a two-form entry primary school for the design task. Subsequently for Experiments Three, Four and Five the brief was adapted and a new site plan drawn up.

The subjects in Experiments Three, Four and Five were asked only to prepare a plan, and not elevations or sections. Neither were they asked for landscaping or other details. The reasons were:

a It was an endeavour to restrict the variety of attributes which they would feel it necessary to specify the important aspects of the school.

b It was believed advantageous to the non-architects who might have been at a considerable disadvantage if they lacked knowledge of draughting skills and conventions.

c In evaluation it again restricted the variety of possible attributes which could be used to make judgements of the plans.

In Experiment Two the design task was for a two-man coast guard station for mass production in glass reinforced polyester. For the students the experiment served the dual functions of being both a learning exercise about grp technology as well as an exercise in design method, in which many of the experimental results were discussed with the subjects at the conclusion.

3.7 Judgement Analysis

In conjunction with the design exercises, in Experiments Two, Three, Four and Five, the designers performed a judgement analysis exercise. This was intended to provide information about their values or priorities, and thus to afford explanations of their balancing of priorities during the design process and therefore in their design solutions. The judgement analysis exercise necessitated the elicitation of the designers' priorities, and the scaling of paired comparisons of them, using a technique recently devised by Saaty (1977). The same technique was also used in the evaluation of alternative designs. Although all the experiments had many underlying similarities, and were all based on testing the theory proposed in chapter 2, their individual organisation was not identical.

In Experiment One attributes used in evaluation of the existing school plans were elicited from, and scaled by, each subject individually.

In Experiment Two a set of four attributes, common to all the subjects, was used. The four were decided by two tutors and the experimenter. The number of attributes was influenced strongly by the whole organisation of the experiment. As a means of experimental control, some subjects scaled attributes both before and after design, others only after design.

In Experiments Three and Four the group of six subjects used 'brain-storming' (Osborn, 1957) to elicit a range of potential attributes. Through discussion, and in Experiment Three an unsuccessful attempt to use a voting procedure based on 'Delphi' (Blohm & Steinbuch, 1973; Wills & Wilson, 1972), a set of six attributes, common to all the subjects, was agreed. These attributes were scaled by each subject both before and after design, and were the attributes with respect to which the design solutions were evaluated. The subjects were also given the opportunity to express other attributes which they had used in comparing the designs.

In Experiment Five the elicitation of attributes took place with each subject separately in order to find the priorities of each individual architect. These were scaled before design. After design each subject had the opportunity of changing the attributes elicited before design, although no subject in the event did so. The attributes were scaled again after design. In evaluation the attributes each subject used were elicited by his comparing the design solutions, and these were the attributes with respect to which the designs were evaluated.

3.8 Redrawing the Designs

In Experiments One, Three, Four and Five the plans were always redrawn to a common scale and format before being evaluated. This was believed to be essential in an experiment of this kind. It was done by Lowe (1972) in his experiments on evaluation, and recommended, though not done, by Cakin (1976). Redrawing has the advantages that:

a The sizes of the plans are readily comparable.

b The orientations of the plans are readily comparable.

c Individual presentation and draughting styles have no effect on the evaluations.

d The amount of information conveyed by each is the same; Porat & Haas (1969) and Cakin (1976) report that quantity of information can have significant effects on decision making.

The disadvantages are that:

a The quantity of information given by the plans can only be as much as that provided by the original plan with the least information (for example, if everyone provided landscape proposals, all the re-drawings could include them, but if one leaves them off they are left off all the redrawn plans).

b The style of redrawing is dependent on who does the redrawing, and may therefore convey some of his prejudices, unintended by the original designer.

c Linked to both previous points, the relative crudeness of the redrawing may eliminate some of the subtlety intended by the designer.

On balance, redrawing was considered essential. The drawing style chosen and used by the experimenter was dictated in part by the time-table. Experiments Three and Four, which each took place in little more than one day, demanded rapid redrawing of the designs, and therefore a rather simple style was adopted. For uniformity, the same style was retained in Experiment Five. In Experiment One previously redrawn plans were adopted. In Experiment Two there were too many plans to make re-drawing feasible. The lack of re-drawing may be significant, as noted in chapter 6.

3.9 Evaluation of Alternative Designs

The method developed by Saaty (1977) for evaluating by means of scaling paired comparisons was used extensively during the evaluation. Its advantages and description are given in the next chapter. In Experiment Two, because of the large number of designs to be evaluated, it was not possible to use paired comparisons and a marking scale 1 - 20 was adopted. It was specified precisely in order to attempt to achieve a degree of comparability between the evaluation marks. In each experiment the subjects themselves evaluated the designs pertaining to, or prepared during, that experiment.

In Experiment One the redrawn school plans were evaluated for overall merit and with respect to each subject's elicited attributes. For overall merit the method of scaling paired comparisons (see chapter 4) was used, and this was followed by the plans being ranked with respect to the subject's attributes.

In Experiment Two the drawings were marked on a 1 - 20 scale, where divisions were precisely specified. The subjects were divided into four groups and members of each group marked the schemes with respect to one attribute. Three tutors also marked the schemes for overall merit using the same 1 - 20 scale.

In Experiments Three and Four the subjects evaluated the redrawn school plans with respect to each of the six common attributes, and for overall merit, using the method of scaling paired comparisons.

In Experiment Five each subject evaluated the redrawn school plans with respect to the attributes which were elicited from him as part of the evaluation process. Each subject also rated the plans for overall merit. Throughout the evaluation the method of scaling paired comparisons was used.

CHAPTER 4 EXPERIMENTAL TECHNIQUES

4.1 Introduction

In chapter 2 the need was established for analysing designers' judgement in conjunction with the design process in order to compare their values with their designs, and to explain their preferences among alternative designs. This chapter describes the criteria for the choice of judgement analysis technique and the details of the methods chosen. It describes how the data are elicited and processed to extract the underlying structure of the judgements. An account is also given of the statistical methods used to compare designers' judgements. Finally an example of one subject's results is given to illustrate the techniques.

4.2 Criteria for Choice of Experimental Techniques

Criteria governing the choice of experimental techniques were established with reference to the author's first hand experience in monitoring design teams, and to existing studies of the design process, particularly in the Department of Design Research. Here Cornforth (1976) has been notably candid in reporting practical difficulties in his experiments to monitor designers.

4.2.1 Input data

a Input data must be relatively easy to gather for the experimenter. If they are not, there is a considerable restriction on the number of subjects who can be monitored. The use of verbal protocols for example (while having other advantages) entails such a restriction, as clear from the limited number of subjects who participate in these experiments. Cornforth (1976), amongst others, is explicit in noting the large effort involved in transcribing and analysing verbal protocols.

The author had already had first hand experience in monitoring two three-day IDE's and more than four weeks of live-project design team observation undertaken as part of the assessment of CEDAR3 computer-aided building design system as described in 1.2.3. During this experience it was found that the subjects' quantitative judgements were frequently made implicity, by extrapolation or interpolation of the computed evaluations, in a way which was not overtly systematic. Hence to obtain measures of the benefits of using the system it was necessary to rely on subjective probability estimates elicited from the subjects. Qualitative judgements were even less possible to monitor than quantitative judgements. These findings emphasised the need for eliciting data from the subjects in a more precise way than passive observation could provide, analogous to the eliciting of subjective probability estimates of the benefits of using a computer aided building design system.

b Eliciting input data must not be too taxing for the subject. If the extracting of data is too taxing the subject may not be able to concentrate sufficiently to give meaningful results, or even to give any results. Again Cornforth (1976) reports openly that his experiments had to be curtailed because of the mental effort involved in completing a repertory grid test at the end of an IDE.

c The input data must be capable of handling both quantitative and qualitative data. Aish (1974) reports the use of connectivity analysis having chosen both the design problem (a console layout) and the data reduction model (connectivity analysis) specifically to eliminate subjective value judgements.

d The data gathered should be of a form which can be readily assimilated by the subject.

In this thesis the attempt has been made to obtain data in a way that is not excessively time consuming, does not tax the subject unduly, relates both quantitative and qualitative aspects of the problem, is in a form with which he is conversant and so may be readily assimilated.

4.2.2 The method of computation

The method of computation must be relatively straightforward. Cornforth (1976) reports difficulty in getting a 'multi-dimensional scaling' (MDS) program to run. In contrast the proposed method of computation is based on the mathematical manipulation of a matrix, which is already available on a pocket calculator, although this was not in fact used.

4.2.3 The output

Criteria governing the output are based on those cited by Tshudi (1972).

a Parsimony: the output should be more simple than the data input.

b Reconstruction: from the output more or less complete recovery of the input should be possible.

c Purification: the output should give a truer, more purified description and thus be said to d reveal latent structure, and

e give information on psychological processes.

Cornforth (1976) in his experiments using an MDS program noted that he as an experimenter had "some difficulty" in interpreting the spatial model output, and that there was limited success when his subjects were involved in the task of interpretation. The presentation of output in the present experiments, as shown for example in figure 5.1, is intended to be readily understood not only by the experimenter but also by the subjects.

4.3 Elicitation of Attributes

Two types of stimuli were used to elicit attributes during the experiments. In each case the principle was that the attributes were elicited from the subjects; they are the attributes which the subjects themselves offer and use. This principle has been followed exactly in Experiments One and Five where each individual subject's attributes are used by him. In Experiments Three and Four the group of six subjects used brainstorming (Osborn, 1957) to elicit attributes, and then agreed a set of six common attributes for the purpose of the experiment. In Experiment Two, the principle was modified, and the attributes were decided by two tutors and the experimenter.

The first means of eliciting attributes comprised giving the subjects a site plan and a brief, and reading a statement such as the following:

> "Consider the implications of planning a two-form entry primary school. What important attributes or qualities would you take into account in planning the school?"

Having been read this statement, in Experiment Five each subject was asked to write down the most important six; in Experiments Three and Four the group of subjects proposed attributes in a brainstorming session. After the brainstorming session had elicited a large number of attributes, a discussion took place at which a set of six common attributes was agreed. An attempt was made to use a voting procedure based on 'Delphi' (Blohm & Steinbuch, 1973; Wills & Wilson, 1972) to agree the six attributes, as described in 7.3, but was abandoned because of overlaps among attributes. These overlaps necessitated considerable discussion, which the voting procedure had tended to eliminate.

The number of attributes, six, was initially decided for Experiment One. There were two reasons. First was the widely recognised observation that the human mind is limited to 7 ± 2 factors for comparison at the same time (Miller, 1956). Second the choice of six rather than seven reduced considerably the task of scaling attributes. The success in which this number resulted in Experiment One led to its continued adoption in Experiments Three, Four and Five. As a check on the validity of this decision, the subjects did have to opportunity of expressing up to twenty attributes in the evaluation phases. An exception to the use of six attributes is Experiment Two where, owing to the organisation of the experiment, the number of attributes was decided as four.

The second means of eliciting attributes was the method of triadic comparisons. Each subject is shown all possible combinations of three items (school plans) from the set of stimuli, and asked to make similarity judgements of them. Given n items he will judge ${}^{n}C_{3}$ triads. For each triad the subject is asked to separate out a pair that share some common and important attribute, which makes them similar and which differentiates them from the third plan. The subject is asked to describe the attribute briefly. For all triads subjects were encouraged to look at alternative ways of pairing off two items before making a final decision. The order in which the triads were given prevented pairs of items appearing in successive triads. Appendices 4.5 and 4.6 show the forms which were completed by the subjects. The maximum number of attributes which could be elicited in this way equals the number of triadic comparisons, that is, when there are six items, ${}^{6}C_{3}$ or twenty. In fact the average number of attributes recorded in these triadic comparisons was between six and seven. This finding vindicated the decision, when it was necessary to specify the number of attributes required, to ask for six.

Both methods of presenting stimuli to elicit attributes have been used successfully in the Department of Design Research. Stansall (1973) and Cornforth (1976) report their use in obtaining data for repertory grid analysis, and the present means of elicitation owe much to their precedents.

It is worthy of note that the attributes elicited in this way may be expressed both positively and negatively; the judge may state that the similar pair share some positive attribute lacked by the third or that the third demonstrates some positive attribute lacked by the other two. It is necessary when compiling the list of attributes with respect to which designs are to be evaluated for the attributes to be expressed in a positive way and this generally necessitates discussion and agreement between the subject and the experimenter.

4.4 Scaling Attributes: 'Prioritization'

The experimental technique which provides a means to analyse designers' judgements of the elicited attributes is that described by Saaty (1977). He entitles the technique 'A scaling method for priorities in hierarchical structures'. It is generally referred to in this thesis as 'prioritization'. The technique has many important properties by which it satisfies the criteria established in 4.2, and which therefore make it suitable for analysing designers' judgemental processes. Probably the most important property of prioritization is that it may be used both to analyse and to express subjects' verbally stated priorities, and to analyse and to express judges' evaluations of alternative designs. Thus the same technique may be used in design and in evaluation.

Prioritization is a means of deriving weights for a set of items according to their subjective importance. In conjunction with the design process, for example, the subjects scale the attributes elicited in terms of their relative importance to the success of the design. In evaluation not only are the attributes scaled for their relative importance, but also the plans are scaled for their relative degree of achievement of each attribute. This gives a form of multi-attribute utility analysis, and allows the weights to be combined, by means of an additive utility model, into an index associated with each plan.

The data input for prioritization consist of scaled judgements of paired comparisons of attributes. Comparisons are implied by value judgements (Nowell-Smith, 1954) so the use of comparisons to elicit such judgements seems apposite. Furthermore as Moroney has said,

> "direct comparison between two items is far more sensitive and discerning than actual measurement at a scale of value. We can tackle the problem of multi-dimensional judgements on the basis of paired comparisons rather than straight ranking" (Moroney, 1951).

The method which Moroney goes on to describe involves the judge deciding for each pair of items which is the more important. These judgements are expressed in a binary matrix, from which it is possible to discern an overall ranking, and a coefficient of consistency derived from the number of circular triads of items occurring.

Saaty's method also involves the judge being presented with all possible combinations of two items from the set to be evaluated. Given n items he will judge ${}^{n}C_2$ pairs, i.e. $n(n-1)/2$. For each pair he is asked to use the pre-specified scale 1 to 9 (shown in appendix 4.1) to decide the weighting of each of the pair. If they are of equal importance this is denoted by each being given the weighting 1; if one is more important it is allocated a weighting on the remainder of the scale, i.e. 2 to 9 according to the degree of importance, and the other of the pair is allocated the reciprocal of that weighting. The logic of this system is that if x is judged more important than y at point 3 of the scale, then y is less important than x to the value of $\frac{1}{3}$.

The judgements are entered in a matrix of the following type:

$$\begin{array}{c|cccc}
 & A_1 & A_2 & \cdots\cdots & A_n \\
\hline
A_1 & w_1/w_1 & w_1/w_2 & \cdots\cdots & w_1/w_n \\
A_2 & w_2/w_1 & w_2/w_2 & \cdots\cdots & w_2/w_n \\
\vdots & & & & \\
A_n & w_n/w_1 & w_n/w_2 & \cdots\cdots & w_n/w_n
\end{array}$$

There is a set of n items denoted by $A_1 \ldots\ldots A_n$. Their weights are denoted by $w_1 \ldots\ldots w_n$. Entries in the matrix are the scaled paired comparisons as elicited from the subject. Each scaled judgement entails two entries: the weight w_i/w_j specified by the judge according to his preference of A_i over A_j, and its reciprocal w_j/w_i.

In practice the convention adopted here is that for a pair A_i, A_j, where i designates a row and j a column, if A_i is judged more important than A_j then the scaled integer is entered as the result in the <u>i</u>th row under the <u>j</u>th column. But if A_j is judged more important than A_i then the reciprocal of the scaled value is entered.

Saaty has taken advantage of the special properties of such a reciprocal matrix to show that it expresses underlying properties of the judgements. First he has shown that the maximum eigenvalue of the matrix gives a measure of the consistency of the judgements forming the entries. For perfect consistency the maximum eigenvalue equals the number of entries. Second he has shown that the normalised eigenvector associated with the maximum eigenvalue of the matrix gives a measure of the importance rating for each element implied by the judge's decisions. If perfect consistency among the judgements is assumed then the relative weights ascribed to items are given by normalising the sums of each row, or alternatively by normalising the entries in any one column.

The scaled judgements need be neither 'cardinally' consistent, nor 'ordinally' consistent. Cardinal consistency would not be expected because judgements do not conform to a precise formula. An example of cardinal consistency would be for A_1 to be related more important than A_3 to the value 9 and A_2 to be relatively more important than A_3 to the value of 3, implying that A_1 must be relatively more important than A_2 to the value $9/3 = 3$. Ordinal consistency expresses the transitivity of preference: if A_1 is relatively more important than A_2 and A_2 is relatively more important than A_3, then A_1 should be relatively more important than A_3. Both types of inconsistency are admitted by the method, and the maximum eigenvalue (λmax) provides a measure of the degree of consistency among the judgements. Saaty notes that there is no apparent relationship between this measure and the coefficient of consistency derived from circular triads in ordinal paired comparisons. However he does present a qualitative statistical test of consistency. For good consistency $(\mu/2)^{\frac{1}{2}} < 1$, where $\mu = (\lambda max - n)/(n - 1)$.

Through a number of trials using the technique Saaty has compared the subjective estimations of relative distances between capital cities, relative brightnesses of illuminated objects, and relative masses of objects, and in each case has been able to show that within certain tolerance limits the normalised eigenvectors corresponded to the actual (normalised) measurements. As these trials confirm the accuracy with which this technique may be used to obtain good subjective estimates of objective facts, this ability would seem to make it valid for use to obtain subjective estimates of subjective values.

Prioritization of attributes elicited from the designer in conjunction with the design process provides two kinds of information about his judgemental process. The maximum eigenvalue provides a measure of the underlying consistency with which the judgements are made. The normalised eigenvectors represent the relative weights of attributes implied by the judgements. The weights are measures of the extent to which the designer will try to achieve each verbally stated attribute in the design. They express his priorities and indicate the trade-offs he is likely to make during the design process. By asking designers to perform prioritization of attributes before and after the design process, the two measures should reveal the effect of the design process on the rating of attributes and on the consistency with which they are scaled.

The same method of scaling may be used to evaluate designs. In evaluation the situation faced by the evaluator is that:

a The designs exhibit many attributes.

b The attributes vary in the values which evaluators ascribe to them.

c The alternatives demonstrate different degrees of fulfilment of the attributes.

Evaluation therefore is a process of eliciting and weighting attributes, weighting the plans with respect to each attribue, and combining these weighted partial judgements into an overall evaluation of all the alternatives. Means of eliciting and ascribing weights to attributes have already been described. Saaty's prioritization may also be used to assign weights to alternative designs, either for their overall merit or with respect to separate attributes. Precisely the same procedure as described above is followed: scaled judgements of paired comparisons of designs are entered into a matrix, and through computation measures of consistency and relative weights found.

To obtain overall evaluations an additive utility model is used. Taking each attribute in turn the relative weight of each design with respect to that attribute is multiplied by the relative weight (relative importance to overall performance) of that attribute. When this has been done for each attribute the products assiciated with each design are added to give an index. The index thus represents the relative value of each design with respect to all the attributes, where the relative importance of the attributes to overall value has also been taken into account. In mathematical terms

$$I = \sum_{n=1}^{n} x_n y_n$$

where I is the index of overall relative value, x_n is the weight ascribed to attribute n for its relative importance to overall value and y_n is the weight ascribed to the design for its relative value with respect to attribute n. 'x' and 'y' are sometimes referred to as 'α' and 'β' (Grant, 1976).

To check that the additive utility model indices genuinely reflect subjects' preferences a comparison has been made between each subject's preferences as expressed by these indices, and as obtained by asking him to scale paired comparisons of designs directly for overall merit. Because in calculating the indices, the eigenvectors are always normalised, so the indices sum to unity. The normalised eigenvectors in overall merit judgements also sum to unity. Thus the two can be directly compared by being drawn on a diagram to the same scale. A correlation coefficient can also be calculated between the two sets of results. The example in 4.7 shows these comparisons. To try to avoid confusion in this particular comparison the term overall merit or overall rating is used when the subjects are asked to rate plans directly for overall merit; the term index is reserved for the combination of partial judgements.

These detailed evaluations using paired and triadic comparisons have been used throughout the experiments except for Experiment Two. In

Experiment Two the large number of schemes made the paired comparison technique impossible, and an alternative means of marking the schemes was adopted. This is described in chapter 6.

In addition to scaling attributes and designs, the subjects have always been asked for a simple rank order. Throughout the experiments the ratings given by eigenvectors are those quoted, and those which have been relied upon. On one or two occasions only, where a subject's consistency has been poor, the simple ranks have been referred to; the text notes explicitly these occasions.

As part of the evaluation process it was believed desirable that each subject should rate his own design scheme. The advantage of paired comparisons is that, because each pair is judged independently, all those judgements relating to the designer's own scheme can later, if required, be eliminated. The results do indicate a general tendency for each designer to rate his own scheme highly, a not unexpected result. However this tendency is not considered to be detrimental to the experimental results, and no attempt has been made to eliminate each designer's evaluation of his own scheme.

The calculations of maximum eigenvalues, normalised eigenvectors, indices and hierarchical clusters (described below) were performed by computer. A program was written by the present author (incorporating a program written by Dr. Mallen: see Declaration) in BASIC to run on the Royal College of Art's Altair mini-computer (appendix 4.7).

4.5 Hierarchical Cluster Analysis

Hierarchical cluster analysis is a technique developed to help identify groupings or clusterings of items inherent in subjective judgements. The technique enables items to be classified into optimally homogeneous groups, that is, objects judged similar are assigned to different groups. Johnson (1967) describes a procedure which constructs a hierarchical system of clustering representations ranging from one in which each of the n objects is represented as a separate cluster to one in which all n objects are grouped together as a single cluster. The result is an explicit and intuitive description of the clustering inherent in the subjective evaluations of designs.

The data for hierarchical cluster analysis consists solely of the n(n-1)/2 similarity measures among the n objects under study. The similarity measures may be obtained in several ways including, for example, obtaining for every pair of objects a subjective rating of similarity. However they may also be obtained from triadic comparisons, the same triadic comparisons in which similarity judgements have been used to elicit attributes, as described in 4.3. The advantages of using these similarity judgements are several. First the triadic comparisons are thereby made to provide data for two parts of the experiment at the same time. Second it obviates the need for another numerical scale to assess similarity between plans. This might prove a difficult task for school plans, and besides a numerical scale is being used to assess merit.

The order in which the subject performs triadic comparisons has been chosen so that, as noted above in 4.3, no pair of items appears in successive triads. This is to minimise any confounding effect which might result from the successive appearance of pairs of items in the triadic comparisons. The order chosen is shown in appendix 4.5.

The similarity judgements are compiled into a similarity matrix of the following kind:

	P_1	P_2	....	P_n
P_1	-	P_1, P_2	....	P_1, P_n
P_2	P_2, P_1	-	....	P_2, P_n
⋮				
P_n	$P_n . P_1$	P_n, P_2	....	-

There is a set of n items denoted by P_1 ...... P_n. Each entry, P_i, P_j, denotes the number of times that pair is judged similar. The diagonal of the matrix is left blank; and the matrix is symmetrical about the diagonal.

Hierarchical clustering schemes are obtained from the matrix using the method detailed by Michon (1969), which is itself based on an algorithm proposed by Johnson (1967).

The output is graphical and takes the form of a similarity tree or dendrogram. For each similarity level this shows the plans which were judged similar at that level. The maximum number of levels is given by n-2, where all the n items are represented as separate clusters; at level 1 all the items are assigned to the same group. The similarity tree is intended to be assimilable both by the experimenter and by the subject. Section 4.7 describes an actual set of results.

4.6 Statistical Measures

The techniques described so far in this chapter have concentrated on eliciting data on designers' subjective judgements. Commonly accepted statistical measures have been used in order to draw inferences from the data. They have been used both to make comparisons within the data from each subject, for example to measure possible changes in priorities resulting from the design process, and between subjects, such as the level of concordance between their evaluations of designs.

The statistics used are Kendall's coefficient of concordance, Spearman's rank correlation coefficient, and the Mann-Whitney U-test. They are described fully in standard texts, such as Siegel (1956) and Moroney (1951). Significance (α) levels achieved are quoted with the results; one-tailed tests are used where applicable, and an α-level of 0.05 has been taken as significant. A computer program was written in BASIC by the author to calculate Kendall's coefficient of concordance (appendix 4.8).

Generally the statistics have been used conventionally. From the theory expressed in chapter 2, a number of research hypotheses arise. These are stated in the experiments as null hypotheses, in the expectation that they will be rejected. The statistics are used to find the probability of the null hypothesis being true, and if this probability is less than one in twenty, the research hypothesis is accepted as true.

Because of the organisation of the experiments, in addition to the conventional use of these statistical measures, it has been necessary to use them unconventionally. For example, one aspect of the tentative theory being explored is that designers will disagree about the relative importance of attributes, and will disagree about the relative merit of designs. When using Kendall's coefficient of concordance it is only really legitimate to use it to measure concordance; the null hypothesis states that there will be no concordance, and a coefficient larger than that which is significant allows the null hypothesis to be rejected. However the non-rejection of the null hypothesis, if the coefficient is below the required level of significance, does not legitimately allow the conclusion that there is significant discordance between the judgements. Nevertheless it has been necessary in these experiments to use Kendall's coefficient of concordance quite extensively. For the most part its use is in the normal and legitimate convention, and the null hypothesis has been stated in the conventional way. But there are occasions when although the null hypothesis has been stated in the conventional way, the expectation is its non-rejection. Where this is the case it has been clearly stated in the drawing of conclusions.

Spearman's rank correlation coefficient is used extensively, and for the most part conventionally. But it is used in an attempt to measure changes in the rating of priorities by each designer. As with Kendall's coefficient the null hypotheses are stated in the conventional way. But the statistic is in fact used in an unconventional way. What has been attempted is to find some measure to decide whether or not there is a significant difference in the rating of attributes before and after design. Spearman's coefficient has been calculated between the before-design and after-design ranks. If the coefficient shows that the correlation is statistically significant it has been assumed that there is no real difference between the two sets of ranks and thus that there has been no change. If the coefficient shows that there is not significant correlation it has been assumed that a definite change in the rating of attributes has taken place. Thus in some cases the non-rejection of the null hypothesis is the expected result, and where this is so it has been stated in the conclusions.

A further unconventional aspect of the experiments is the performing of so many tasks by the subjects. It is much more usual to hold constant all the possible factors except the one or two in which the experimenter is interested and then to vary these under controlled conditions. Such control however is extremely limiting in experiments on the design process. It seems reasonable on the other hand to classify the present study not only as the testing of a theory but also as an exploration and a search for underlying patterns in the design process, as they can be recorded in design studio conditions. Part of the experiments entail trying a number of tests to identify potential correlations, and it may well be the case that the indications of potentially fruitful avenues discovered in these explorations will prompt laboratory based confirmations, given the present experimenter's reservations about excessively controlled experimental conditions, mentioned in 3.3.

In strong support of the present experiments it is claimed that, despite the occasionally unconventional use of statistical measures, and the rather large number of variables, these experiments, which have developed from the passive monitoring of design processes, by stating and testing hypotheses precisely and in some detail, present an important and valuable departure from that approach.

4.7 An Example

This section provides an example to illustrate in detail the techniques. The data from S5.6's evaluation of school plans will be used. The full experiment is described in chapter 9.

First through triadic comparisons of school plans, attributes for evaluation were elicited and similarity judgements of plans were obtained. In these triadic comparisons the subject completed the form (appendix 4.5) as shown below. The plans are represented by letters A to F, and are shown in figure 9.2.

Triad			Attribute
A	B	C	Orientation
A	D	E	Clear geometry
B	D	F	Clear geometry
B	C	E	Orientation
A	E	F	Compactness
A	C	D	Compactness
B	E	F	Symmetry

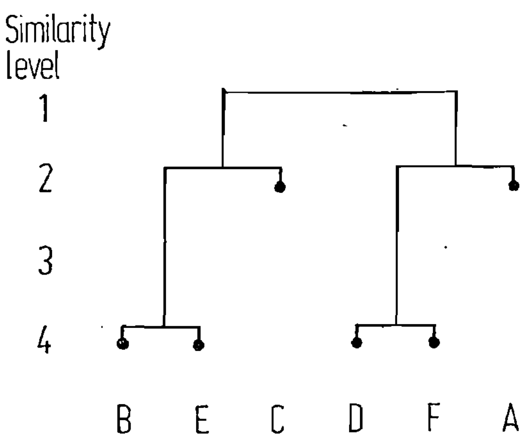
Triad			Attribute
A	C	F	Symmetry
C	D	E	Symmetry
A	B	D	Symmetry
C	E	F	Orientation
B	C	D	Symmetry
A	D	F	Compactness
A	B	E	Symmetry
C	D	F	Orientation
A	B	F	Symmetry
B	D	E	Clear geometry
A	C	E	Orientation
B	C	F	Symmetry
D	E	F	Symmetry

The similarity judgements were entered into a matrix thus:

	A	B	C	D	E	F
A	-	1	0	2	1	2
B	1	-	2	0	4	0
C	0	2	-	1	2	0
D	2	0	1	-	0	4
E	1	4	2	0		1
F	2	0	0	4	1	

	A	BE	C	DF
A	-	1	0	2
BE	1	-	2	0
C	0	2		1
DF	2	0	1	

	ADF	BEC
ADF	-	1
BEC	1	-



This matrix is searched for the most similar items. Rows and columns are collapsed by taking the highest values of pairs of cells, and the process is repeated until the matrix is reduced completely. From the successive contractions can be seen the derivation of the hierarchical clustering shown, both the clusters and the similarity levels.

Second the elicited attributes are scaled. The four attributes elicited were 'orientation' (a), 'compactness' (b), 'symmetry' (c), 'clear geometry' (d). The subject was given the scale to be used (appendix 4.1) and was read the instructions at the top of the form shown in appendix 4.2, slightly modified because there were four attributes not six. He was given the attributes in pairs in the order shown on the form, again modified because there were only four. For each pair he estimated their importance in the planning of a primary school relative to one another. The scaled judgements were recorded on the form by the experimenter, and transfered to a matrix thus:

	a	b	c	d
a	1	4	5	1
b	1/4	1	5	1
c	1/5	1/5	1	1/3
d	1	1	3	1

From this matrix it can be seen that 'a' was more important than 'b' to the value 4, and so on. Through computation the maximum eigenvalue was found to be 4.282 and the normalised eigenvectors

(0.449, 0.218, 0.068, 0.265)

Third the school plans are scaled for their merit with respect to each of the attributes. For each attribute a form of the type reproduced in appendix 4.3 was shown to the subject who was then given all possible pairs of plans in turn. The experimenter recorded the scaled judgements. For each of the attributes the following matrices were recorded.

Attribute		A	B	C	D	E	F
a	A	1	3	7	1/2	2	1/2
	B	1/3	1	7	1/3	1/3	1/4
	C	1/7	1/7	1	1/7	1/7	1/8
	D	2	3	7	1	3	1
	E	1/2	3	7	1/3	1	1/3
	F	2	4	8	1	3	1
b	A	1	1/5	1/5	1/5	1/4	1/6
	B	5	1	1/5	1/5	1/3	1/4
	C	5	5	1	3	4	1/2
	D	5	5	1/3	1	3	1/3
	E	4	3	1/4	1/3	1	1/4
	F	6	4	2	3	4	1
c	A	1	5	7	1/3	7	1/3
	B	1/5	1	4	1/8	1	1/7
	C	1/7	1/4	1	1/7	1	1/8
	D	3	8	7	1	7	1
	E	1/7	1	1	1/7	1	1/8
	F	3	7	8	1	8	1
d	A	1	5	1/3	1/3	4	1/4
	B	1/5	1	1/6	1/5	1	1/7
	C	3	6	1	1/3	6	1/3
	D	3	5	3	1	5	1
	E	1/4	1	1/6	1/5	1	1/7
	F	4	7	3	1	7	1

Through computation of the above matrices the normalised eigenvectors were found to be as follows:

	Plans					
Attributes	A	B	C	D	E	F
a	0.184	0.084	0.025	0.279	0.133	0.296
b	0.033	0.067	0.282	0.177	0.097	0.345
c	0.186	0.055	0.032	0.343	0.038	0.345
d	0.108	0.037	0.184	0.296	0.038	0.337

The index for plan A is given by

$$I = \sum_{n=1}^{n=4} x_n y_n \qquad \text{(see } \underline{4.4}\text{)}$$

$$= (.449 \times .184) + (.218 \times .033) + (.068 \times .186) + (.265 \times .108)$$

$$= .083 + .007 + .013 + .029$$

$$= .131$$

This computation for each plan gives the indices listed in table 9.8 for subject S5.6.

Additionally the subject is asked to scale paired comparisons of school plans for their overall merit as schools, using the form shown in appendix 4.4. The results are given for subject S5.6 in table 9.9.

Figure 9.1 for S5.6 shows his results graphically. The similarity tree described above is given, showing the subjective clusterings of school plans derived from the similarity judgements. Below this, the additive utility model indices (dotted line) and the eigenvectors given in overall evaluation (solid line) are plotted to the same scale. In this way the figure records, in easily assimilable form, the comparison between these three methods of evaluation. As figure 9.1 shows for subject S5.6, with the slight exception of the weights ascribed to plans A and F, there is a high level of correlation between the indices and the overall evaluation eigenvectors. This is checked by calculating Spearman's rank correlation coefficient between the two sets of results. Furthermore there is an obvious correspondence between these two sets of results and the hierarchical clustering: the members of cluster B-E-C are rated low, those of cluster D-F are rated high. Only the overall merit rating of plan A is slightly lower than might be expected. Tabulations of these comparisons for all subjects are given in the experiments.

4.8 Summary

The techniques which have been chosen have important features indicative of their suitability in the circumstances of the present thesis.

1 The attributes which form the subject matter of the judgements are those elicited from the subjects themselves. They are the attributes which the designer himself expresses verbally and there is no reason to

believe that designers should be unable or unwilling to describe the attributes they value. The means of eliciting the attributes is not believed to be excessively time consuming, nor exhausting. Attributes may be both quantitative and qualitative.

2 The judgement phase of weighting the attributes entails comparative judgements of pairs of attributes. It is a process which is intuitively reasonable and not onerous for the subject.

3 Computation from the input is a straightforward matter of calculating from the matrix of pairwise comparisons the maximum eigenvalue, and its associated eigenvector.

4 Output is readily assimilable by the experimenter and by the subjects. The notion of relative weighting of items, as given by the normalised eigenvector, is widely recognised, and the measure of consistency given by the maximum eigenvalue may be easily explained.

5 In the evaluation phase the above four points apply. Additionally the triadic comparisons are an efficient means of eliciting data and are intuitively reasonable. The resulting cluster analysis is simple to compute and the output easily comprehended. The process of scaling paired comparisons of plans with respect to the various attributes is the most involved part of eliciting the data, and may entail one hundred or more paired comparisons by each subject. In the context of these experiments the subjects have performed this number of comparisons without complaint. If perfect consistency is assumed only one row of the matrix would need to be completed. This would reduce dramatically the number of paired comparisons from, for instance, over one hundred to about thirty and would simplify somewhat the ease of using the technique. The index which is output gives a measure of the overall weighting of the plans as a combination of partial judgements, and is almost as easy to assimilate as simple ranking.

As an experimental check on the answers, the subjects also scale judgements of school plans for their overall merit. In this way the same stimulus items, the school plans, have in fact been evaluated by three different methods: similarity judgements, partial judgements and overall merit. By comparing the results given by these methods, conclusions may be drawn about their ability to represent reliably subjective evaluations. The degree to which all three sets of results correspond provides an indication of the confidence which may be placed in the techniques and in their ability to provide meaningful and useful results. The final justification will lie in whether or not what Bross (1953) has called the pragmatic principle is satisfied: does it work?

CHAPTER 5 EXPERIMENT ONE

5.1 Introduction

The primary aim of the first experiment was to try out and to develop experience of the proposed experimental techniques: Saaty's 'prioritization', the use of triadic comparisons to elicit attributes, and hierarchical cluster analysis. Additionally it was hoped to find out whether lack of architectural training would prevent meaningful results from being obtained with these techniques.

Six subjects took part. S1.1 and S1.2 were architects of more than ten years experience in architectural practice, S1.3 and S1.4 were students with first degrees in architecture, S1.5 and S1.6 were non-architects, one an ergonomist, one a secondary school teacher. Each subject was interviewed individually by the experimenter.

Attributes for the design of a school plan were elicited from the subjects, and rated using scaled paired comparisons. The subjects were then shown the existing school plans, and the method of triadic comparisons was used to elicit attributes for evaluation. These attributes were also rated using scaled paired comparisons. The school plans were rated using scaled paired comparisons.

The designs being evaluated were alternative sketch plans for a two-form entry primary school in Hertfordshire, which had been designed in the Local Authority offices. Each plan was drawn out on a 6" x 4" card in a standard format using felt-tip pen. Rooms were labelled and access entrances shown. An outline brief for the building was also given.

5.2 Hypotheses

The experiment was designed to test a number of hypotheses. All are expressed as null hypotheses. Additionally the experimental techniques were evaluated for their ability to provide meaningful results.

5.2.1 Consistency and architectural experience

Hypothesis 1 That there would be no significant correlation between the internal consistency achieved in scaling paired comparisons of attributes for design and the degree of architectural experience.

Hypothesis 2 That there would be no significant correlation between the internal consistency achieved in scaling paired comparisons of school plans and the degree of architectural experience.

5.2.2 Numbers of attributes used in evaluation, and architectural experience

Hypothesis 3 That there would be no significant correlation between the number of attributes used in evaluation and the degree of architectural experience.

5.2.3 Concordance between evaluations of school plans

Hypothesis 4 That there would be no significant concordance between the subjects' overall ratings of school plans.

5.2.4 Overall rating of school plans, hierarchical cluster analysis and the use of an 'index'

Hypothesis 5 That for each subject there would be no significant difference between the overall rating of school plans and the hierarchical clustering of school plans.

Finally an attempt was made to obtain for each subject an index for the school plans derived from the rating of attributes and the rating of school plans with respect to attributes. The index for each plan was compared with the overall rating of each plan.

5.3 Experimental Method

In order to test the hypotheses the experiment was organised as follows:

1 Each of the six subjects (S1.1, S1.2, S1.3, S1.4, S1.5, S1.6) was shown an outline brief of a two-form entry primary school. Each was asked to write down the six attributes he considered to be important in the planning of a primary school to satisfy the brief given. Each attribute was copied onto a card. Each subject was shown the numerical scale to be used in scaling paired comparisons of attributes and told how to use it. All possible fifteen pairs of attributes were shown in turn

to the subject who rated them, using prioritization, in terms of their relative importance in the planning of a primary school. Each subject was also asked to give a simple rank order of his six attributes for design.

2 Each subject was shown the six existing school plans (figure 5.2). These were withdrawn and shown again to him in threes. In these triadic comparisons each subject was asked to separate out two of the three by virtue of their sharing a common attribute, from the third which does not demonstrate this attribute. The attribute and the similar pair chosen were recorded for all twenty possible triads shown to each subject.

3 Each subject was shown all possible pairs of school plans and for each pair asked to rate numerically how well the better of the pair would function as a school, using prioritization.

4 Each subject used prioritization to rate the attributes he had used in performing triadic comparisons of school plans (i.e. attributes for evaluation), and also gave a simple rank order of these attributes.

5 Taking each of their own attributes for evaluation in turn, each subject ranked the school plans for their performance with respect to that attribute.

Additionally S1.2 used prioritization to rate the plans. He scaled paired comparisons of school plans for their performance with respect to each of his attributes for evaluation.

5.4 Results

5.4.1 Consistency and architectural experience

Hypothesis 1 was tested by calculating Spearman's rank correlation coefficient between the rank order of maximum eigenvalues achieved by each subject in scaling paired comparisons of attributes for design and the rank order of the degrees of architectural experience. Table 5.3 shows the data. The coefficient was found to be 0.271 which was not significant (one-tailed test). The null hypothesis was not rejected.

Hypothesis 2 was tested by calculating Spearman's rank correlation coefficient between the rank order of maximum eigenvalues achieved by each subject in scaling paired comparisons of school plans and the rank order of the degrees of architectural experience. Table 5.3 shows the data. The coefficient was found to be 0.157 which was not significant (one-tailed test). The null hypothesis was not rejected.

5.4.2 Numbers of attributes used in evaluation and architectural experience

Hypothesis 3 was tested by calculating Spearman's rank correlation coefficient between the rank order of the number of attributes used in evaluation and the rank order of the degrees of architectural experience. Table 5.2 shows the data. The coefficient was found to be 0.043. The null hypothesis was not rejected.

5.4.3 Concordance between evaluations of school plans

Hypothesis 4 was tested by calculating Kendall's coefficient of concordance between the six sets of ranks of school plans derived from scaled paired comparisons. Table 5.4 shows the data. The coefficient was found to be 0.368 which was significant ($\alpha = 0.05$). The null hypothesis was rejected.

5.4.4 Overall rating of school plans, hierarchical cluster analysis and the use of an index

Hypothesis 5 was tested by comparing the overall rating of school plans and the hierarchical clustering of school plans. Figure 5.1 shows the data. The following table gives a verbal estimate on a four point scale 'high', 'medium', 'low', 'no' of the correspondence for each subject between the overall rating of school plans and the hierarchical cluster analysis.

Subject	
S1.1	Medium correspondence. The worst plan (F) separated out. The middle three plans (C, E, D) judged similar.
S1.2	High correspondence. The best plan (C) separated out. The worst three plans (B, E, F) judged similar. Second and third plans judged between the first and last three.

S1.3 *Medium* correspondence. The best two plans (C, D) judged similar. The worst three judged similar. Plan (E) added to the best two (note, if E and B, which were rated almost exactly equal, had their ratings reversed then the similarity tree would show the best three and the worst three in two clusters).

S1.4 *No* correspondence.

S1.5 *No* correspondence.

S1.6 *High* correspondence. The best plan (A) separated out. The worst two (C, E) judged similar. The middle three (B, D, F) judged similar.

The null hypothesis was not entirely rejected.

In the case of S1.2 an index for each of the school plans was obtained by combining the partial judgements of attributes and of school plans with respect to attributes using Saaty's technique, as described in Chapter 4. Table 5.5 shows the data. In the other five cases weightings had been obtained for attributes, but only a rank order was available of the school plans judged with respect to each attribute.

In order to obtain an approximate index, weights were attached to the school plans when they were judged with respect to each attribute. These weights were based inversely on the rank order. First the schemes were allocated points in order of merit, six points to the school ranked first, one point to the school ranked sixth. These points were normalised by having their sum equal to unity. This gave, generally, the following weights to the plans in descending order of merit: 0.286, 0.238, 0.190, 0.143, 0.095, 0.048. Tied ranks were taken into account by being given equal points, while at the same time the total number of points (21) remained the same. Broadly the assumption being made is that the plans are weighted with equal distances between them. To obtain the overall index for each plan these 'weights' were multiplied by the weights of each attribute derived from prioritization, and the products added. Table 5.5 shows the use of Saaty's method for subject S1.2 who scaled paired comparisons of attributes and of school plans with respect to each attribute. Table 5.6 shows as an example the approximation used in this experiment to obtain an index for subject S1.3. Table 5.7 shows the approximate indices for subjects S1.1, S1.3, S1.4, S1.5, S1.6 and includes the index for S1.2 using Saaty's method.

The ratings given in tables 5.4 and 5.7 are shown as subjective evaluation profiles in figure 1. The comparison between the two profiles shows to what extent the combination of partial judgements is equivalent to overall preference, that is, whether the present experiment reveals the way in which overall preferences may be accounted for by partial judgements. Although the results are not precisely as expected there are grounds for some confidence.

The analyses of S1.2's judgements are closest to those expected in that there is a close correspondence between the shapes of the two profiles and the hierarchical cluster analysis of the similarity judgements. The two rank orders of plans (overall rank: C D A E B F, rank derived from indices: C A D B E F) are slightly different. But the detailed analyses of judgements help to show underlying similarities. In all three analyses the most preferred plan, (C), is separated out and rated highly. The worst three plans, (B, E, F), are all clustered together and rated similarly. However by comparison with the overall ratings the intermediate plans, (A, D), have their rank orders reversed according to the indices.

For S1.3 the two subjective evaluation profiles have similar shapes and the comparison between the overall rank (table 5.4) and the indices (table 5.7) showed that the most preferred plan, (D), and the least preferred plan, (F), were the same in each case. A similar result was found in the case of S1.5; the two subjective evaluation profiles have similar shapes and the least preferred plan, (A), was the same in each case, while plan C scored highly in each case.

The other subjects' results are less encouraging in this respect.

5.5 Discussion of results

This experiment was meant only to be a pilot study. The small number of subjects and the organisation of the experiment barely warrant the drawing of inferences from the results. Nevertheless the results provide indications of the possible worth of experiments of this type.

The result of testing hypothesis 1 suggests that architectural experience is not essential to performing consistent judgements about attributes for school planning. There was no correlation between consistency and architectural experience. The result of testing hypothesis 2 suggests further that architectural experience is not essential to performing consistent judgements of school plans. From these two results it seems that the techniques being employed, particularly the use of Saaty's prioritization, do not necessitate subjects with architectural training. Similarly in the triadic comparisons even those with no architectural training had no difficulty in making similarity judgements among plans.

The result of testing hypothesis 3 is an important one. The maximum number of attributes for evaluation which any subject could have used was twenty (i.e. the number of triadic comparisons). It may be concluded that no respondent was restricted by the method to fewer attributes than he would otherwise have considered. The average number of attributes used was 6.667, a figure in line with the number of attributes for design initially chosen, and strongly supportive of the use of this number of attributes as described in 4.3.

The result of testing hypothesis 4 shows that there was significant agreement between the ranks (derived from scaled paired comparisons) of the school plans given by the six subjects. As there is agreement the six sets of judgements may be combined to discover the group's overall preferences. This is done by adding the ranks given to each plan in table 5.4. This gives an overall preference for the group: D, B and C equal, A, E, F.

All these six plans are generically similar, particularly in the assembly hall position and classroom arrangements; they give the appearance of having been prepared by one architect (unfortunately the history of the plans is not available). The differences between the plans are differences not of dramatically different value systems in their generation; rather they seem to represent a gradual refinement of the same basic idea. For these reasons when they are evaluated the evaluators are not only evaluating them for the different value systems they manifest, but also for the different levels of skill they exhibit in their execution. This seems the most probable explanation for the significant level of agreement between the judges.

This pilot study was intended essentially to provide experience of using the proposed experimental techniques, and to ascertain their ability to give meaningful results. The methods of using paired and triadic comparisons to elicit and to rate attributes proved to be highly effective. They enable four hypotheses to be tested formally. Comparison of the results of performing hierarchical cluster analysis on the similarity judgements of plans with the weighted rank order derived from paired comparisons of plans was not an unqualified success but was nevertheless encouraging. The attempt to derive an index from the weights of attributes (using Saaty's prioritization) and the weights of plans with respect to each attribute (using Saaty's prioritization) was successful with the one architect who performed the full range of necessary judgements. There was less success with the other subjects when the weighting of plans with respect to each attribute was based only on the ranks, though the results were not totally discouraging.

The experiment was conducted with each subject individually under informal conditions at the experimenter's workspace, at the experimenter's home and in two cases at the subjects' homes. The impact of these varying conditions on the results is not easy to assess. However the maximum eigenvector, or internal consistency measure, is a valuable indicator of the underlying rationality with which the judgements are being made. The fact that some meaningful results have been achieved in this experiment is an encouraging sign and the results are sufficiently worth while for these techniques to be used again with greater control of the experimental conditions.

Table 5.1 Description and rating of attributes for school plan design

Subject	Attribute	Eigenvector	Description of attribute
S1.1	a	.278	Friendly interior environment (colours, furniture, etc)
	b	.205	Natural daylight and ventilation in classrooms and dining room
	c	.156	Flexible use of teaching areas - including hall and dining room
	d	.203	Quiet area for activities disturbed by or producing noise (music, concentrated work, use of video equipment, etc)
	e	.103	Separate access for services (delivery), small through traffic in teaching areas and library
	f	.056	Space for changing/washing before/after physical exercise/external play
S1.2	a	.427	Orderliness
	b	.307	Clear circulation pattern
	c	.083	Daylight
	d	.042	Sensible disposition of service/serviced functions (e.g. kitchen, wc's, etc)
	e	.065	Clear entrance arrangements (people, vehicles)
	f	.077	Sensible orientation
S1.3	a	.132	Internal circulation
	b	.122	Separation of teaching areas and service areas
	c	.264	Relationship of various spaces to the site
	d	.120	Main entrance - relationship to site and rest of school
	e	.137	Flexibility
	f	.225	Organisation and grouping of spaces based on organisation of school

Sub-ject	Attri-bute	Eigen-vector	Description of attribute
S1.4	a	.104	Site geometry/access/orientation
	b	.135	Building cost
	c	.301	Running/maintenance cost - energy/labour costs
	d	.165	System building - probability of reproduction
	e	.090	Degree of isolation of spaces - variation and spatial zoning
	f	.205	Style - historical perspective
S1.5	a	.180	Position of classrooms for ease of access for administration and teaching staff. No rooms to be out on a limb.
	b	.197	Toilet, cloakroom and washing facilities to be adequate and, ideally, near to each classroom.
	c	.104	Administration block and staffroom near main entrance but part of school.
	d	.179	Classrooms should have one window facing south - and should have plenty of window space.
	e	.115	Dining facilities - kitchen and dining room should adjoin and be slightly apart from the main body of the school.
	f	.226	Hall should be central and easily accessible from all rooms.
S1.6	a	.290	Room sizes and facilities to comply with all regulations.
	b	.138	Clear separation of children and staff/service areas for noise and safety reasons.
	c	.019	Exterior of building inviting and friendly.
	d	.028	Views from interior designed to take advantage of views onto playing fields and exterior features.
	e	.100	Interior bright, exciting, cheerful and colourful.
	f	.425	Position and size of windows for daylight, heat loss, glare.

Table 5.2 Description and rating of attributes for evaluation of school plans

Subject	Attribute	Eigenvector	Description of attribute
S1.1	a	.048	Relationship of classroom to library
	b	.107	Dining room - natural light and ventilation
	c	.086	Compactness
	d	.104	Access to external play area
	e	.026	Courtyard position
	f	.194	Orientation of classrooms
	g	.182	Circulation
	h	.109	Position of classrooms
	i	.049	Relationship of entrance to library
	j	.096	Divided classrooms
S1.2	a	.449	Orderliness
	b	.322	Clear circulation
	c	.078	Daylight
	d	.071	Entrance arrangements
	e	.080	Sensible orientation
S1.3	a	.145	Formal organisation and ordering principles
	b	.554	Grouping/dispersal of classes
	c	.233	Access to classes
	d	.067	Service areas grouping
S1.4	a	.138	Location of entrance and delivery
	b	.204	Arrangement of hall/dining room/kitchen
	c	.424	Logic of geometry and zoning
	d	.038	Length of external wall
	e	.107	Entrance planning
	f	.058	Direct access between hall and classrooms
	g	.032	Courtyard size and location

Sub-ject	Attri-bute	Eigen-vector	Description of attribute
S1.5	a	.292	Orientation of classrooms - south facing for sunlight
	b	.074	Relationship of entrance to library
	c	.101	Compactness
	d	.162	Dining room and kitchen away from classrooms
	e	.170	Dining area - daylight and natural ventilation
	f	.030	Orientation of entrance
	g	.041	Entrance separated from service delivery
	h	.129	Integration of administrative offices into main part of school
S1.6	a	.147	Delivery away from classrooms
	b	.108	Delivery away from entrance
	c	.107	Administration near delivery
	d	.240	Orientation of classrooms for sunlight
	e	.347	Administration offices isolated from classroom (for noise)
	f	.051	Administration offices near hall

Table 5.3 Consistency achieved in scaling attributes

Subject	Attributes for design	Attributes for evaluation		School plans
	λmax	Number used	λmax	λmax
S1.1	6.246	10	10.848	6.167
S1.2	6.527	5	5.537	6.976
S1.3	6.638	4	4.338	6.278
S1.4	7.064	7	7.568	7.246
S1.5	6.192	8	9.169	6.779
S1.6	9.114	6	6.647	6.489

Table 5.4 Overall rating of school plans: eigenvectors derived from paired comparisons, corresponding ranks and coefficient of concordance (W)

Plan	Eigenvectors						Ranks						W	Significance level
	S1.1	S1.2	S1.3	S1.4	S1.5	S1.6	S1.1	S1.2	S1.3	S1.4	S1.5	S1.6		
A	.108	.182	.115	.311	.059	.580	5	3	5	1	6	1		
B	.318	.040	.122	.160	.342	.101	1	5	3	3	1	3		
C	.132	.392	.277	.297	.246	.054	4	1	2	2	2	5	0.368	.05
D	.158	.281	.307	.098	.179	.117	3	2	1	4	3	2		
E	.228	.080	.121	.086	.080	.048	2	4	4	5	5	6		
F	.057	.024	.058	.048	.090	.100	6	6	6	6	4	4		

Table 5.5 Rating of attributes and rating of school plans with respect to each attribute, given by subject S1.2

Attribute	Weighting of attributes	Plans A	B	C	D	E	F
a	.449	.366 .164	.120 .054	.309 .139	.115 .052	.052 .023	.038 .017
b	.322	.276 .089	.040 .013	.415 .113	.156 .050	.084 .027	.031 .010
c	.078	.371 .029	.217 .017	.140 .011	.176 .014	.050 .004	.046 .004
d	.071	.095 .007	.167 .012	.286 .020	.167 .012	.238 .017	.048 .003
e	.080	.046 .004	.041 .003	.315 .025	.309 .025	.186 .015	.103 .008
INDEX		.293	.009	.308	.153	.086	.042

Note: The figures in the top left corners of each entry represent the relative weighting (normalised eigenvectors) of the school plans judged with respect to each attribute. The figures in the bottom right corners of each entry represent this relative weighting multiplied by the relative weighting of the attribute. The index is the sum of these products for each plan.

Table 5.6 Rating of attributes and rating of school plans with respect to each attribute, given by subject S1.3

Attribute	Weighting of attribute	Plans					
		A	B	C	D	E	F
a	.145	.238 .035	.143 .021	.286 .041	.190 .028	.095 .014	.048 .007
b	.554	.286 .158	.143 .079	.190 .105	.238 .132	.095 .053	.048 .027
c	.233	.095 .022	.143 .033	.238 .055	.286 .067	.190 .044	.048 .011
d	.067	.286 .019	.143 .010	.238 .016	.190 .013	.095 .006	.048 .003
INDEX		.234	.143	.217	.240	.117	.048

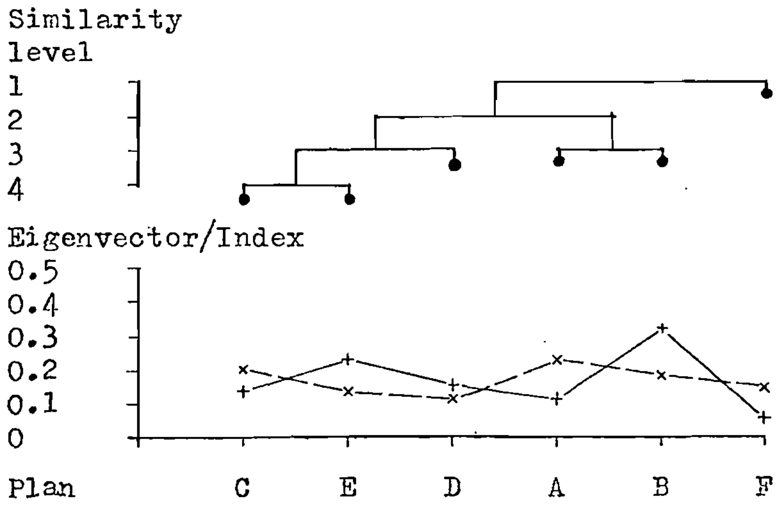
Note: The figures in the top left corners of each entry represent the relative weighting of the school plans derived from their rank order judged with respect to each attribute. The figures in the bottom right corners of each entry represent this relative weighting multiplied by the relative weighting of the attributes. The index is the sum of the products for each plan.

Table 5.7 Overall rating of school plans: indices

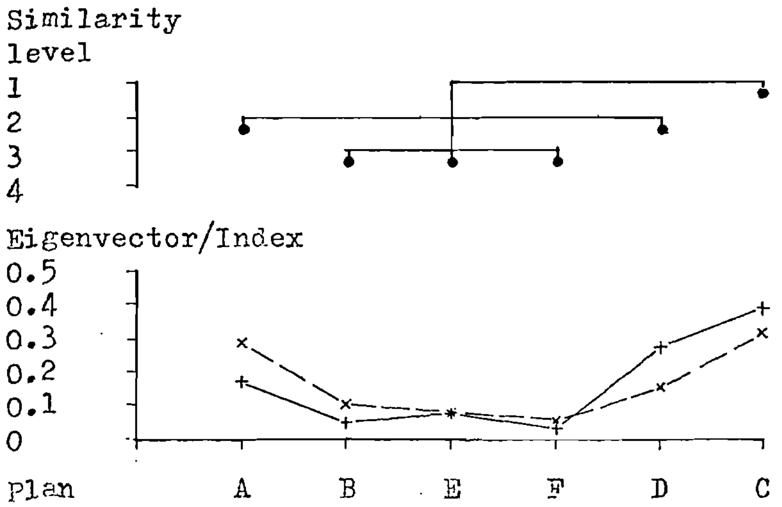
Plan	Subject											
	Indices						Ranks					
	S1.1	S1.2	S1.3	S1.4	S1.5	S1.6	S1.1	S1.2	S1.3	S1.4	S1.5	S1.6
A	.226	.293	.234	.264	.085	.162	1	2	2	1	6	3
B	.187	.099	.143	.127	.170	.114	3	4	4	5	4	5.5
C	.205	.308	.217	.203	.215	.205	2	1	3	2	1	2
D	.112	.153	.240	.202	.204	.248	6	3	1	3	2	1
E	.123	.086	.117	.152	.189	.157	5	5	5	4	3	4
F	.145	.042	.048	.053	.136	.114	4	6	6	6	5	5.5

Figure 5.1 Comparison between each subject's similarity judgements of school plans, the weighting given in overall evaluation and the additive utility model indices.

Subject S1.1



Subject S1.2



Subject S1.3

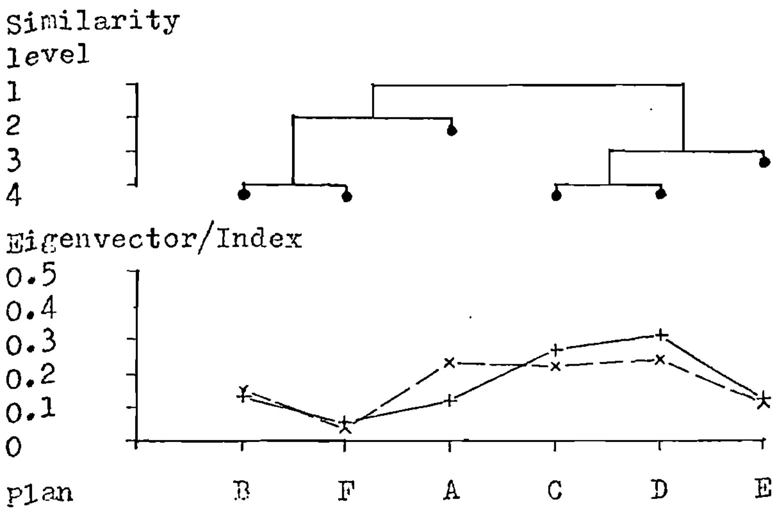
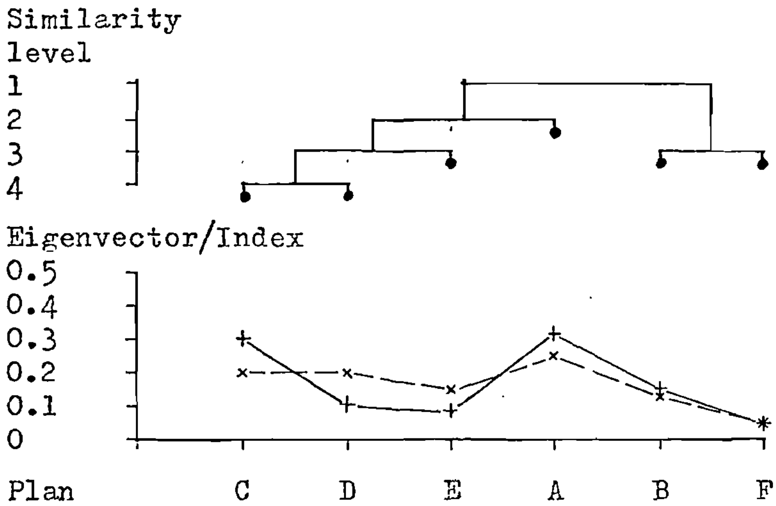
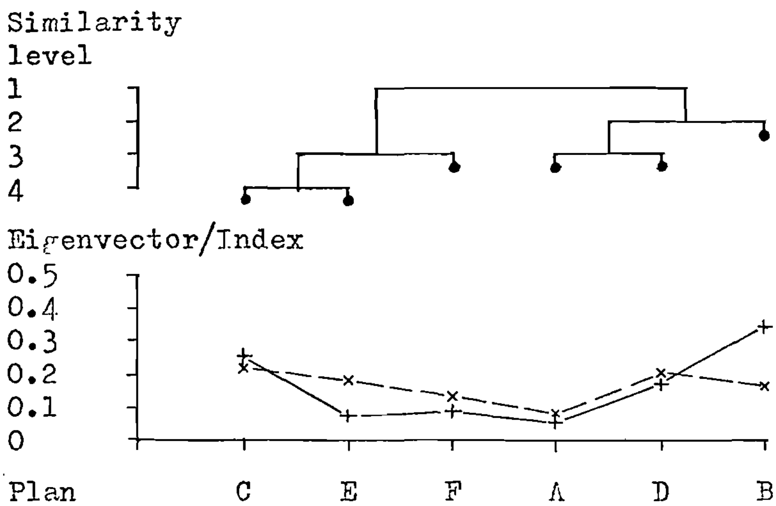


Figure 5.1 continued

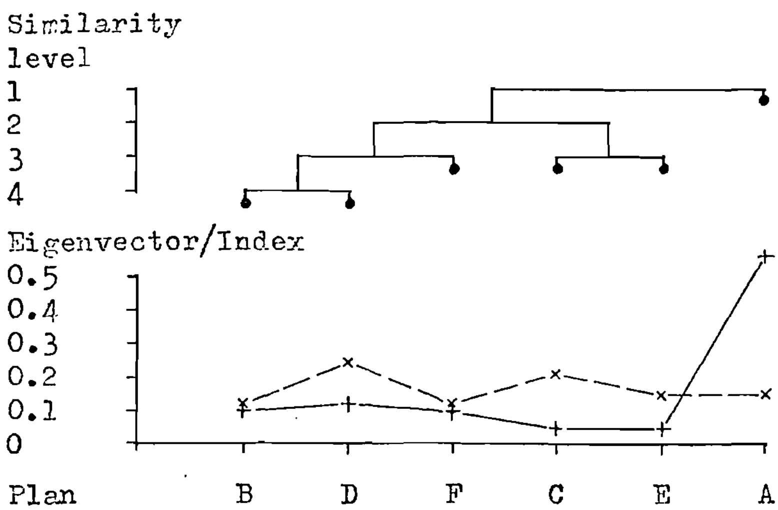
Subject S1.4



Subject S1.5



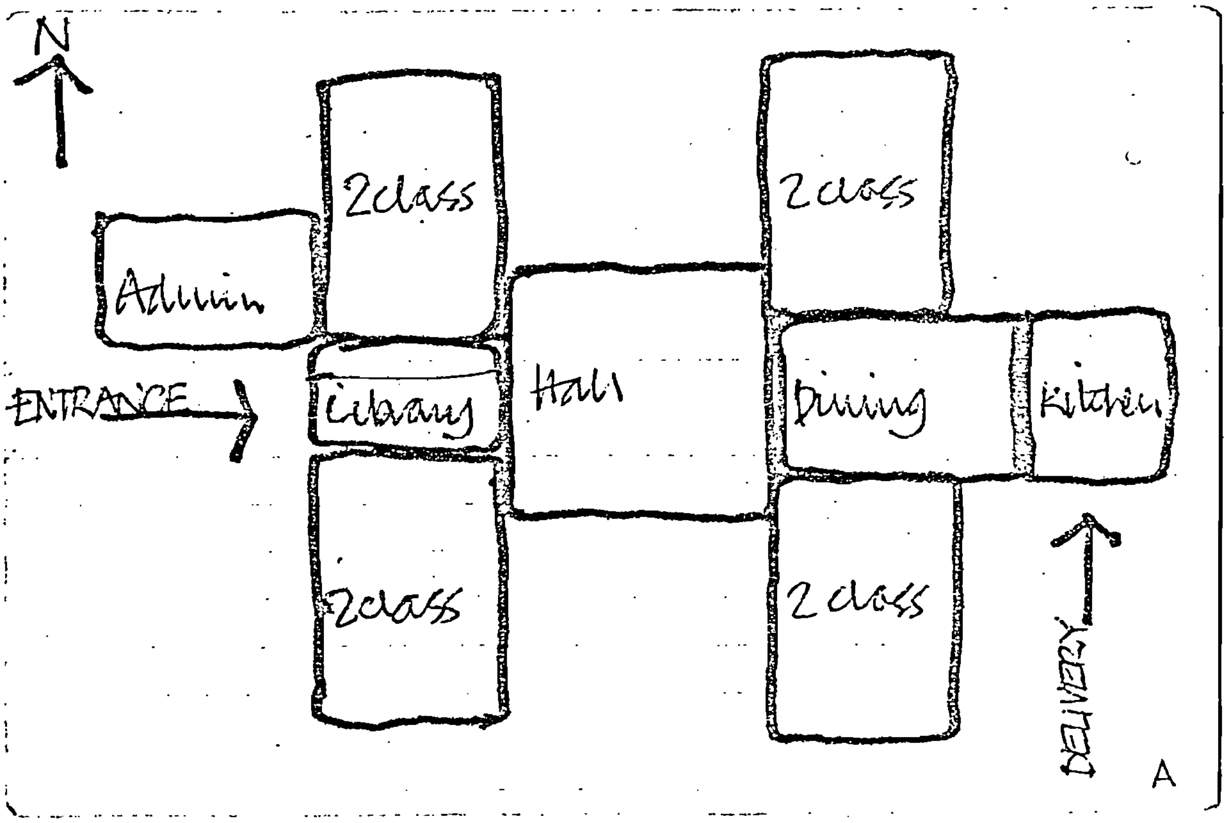
Subject S1.6



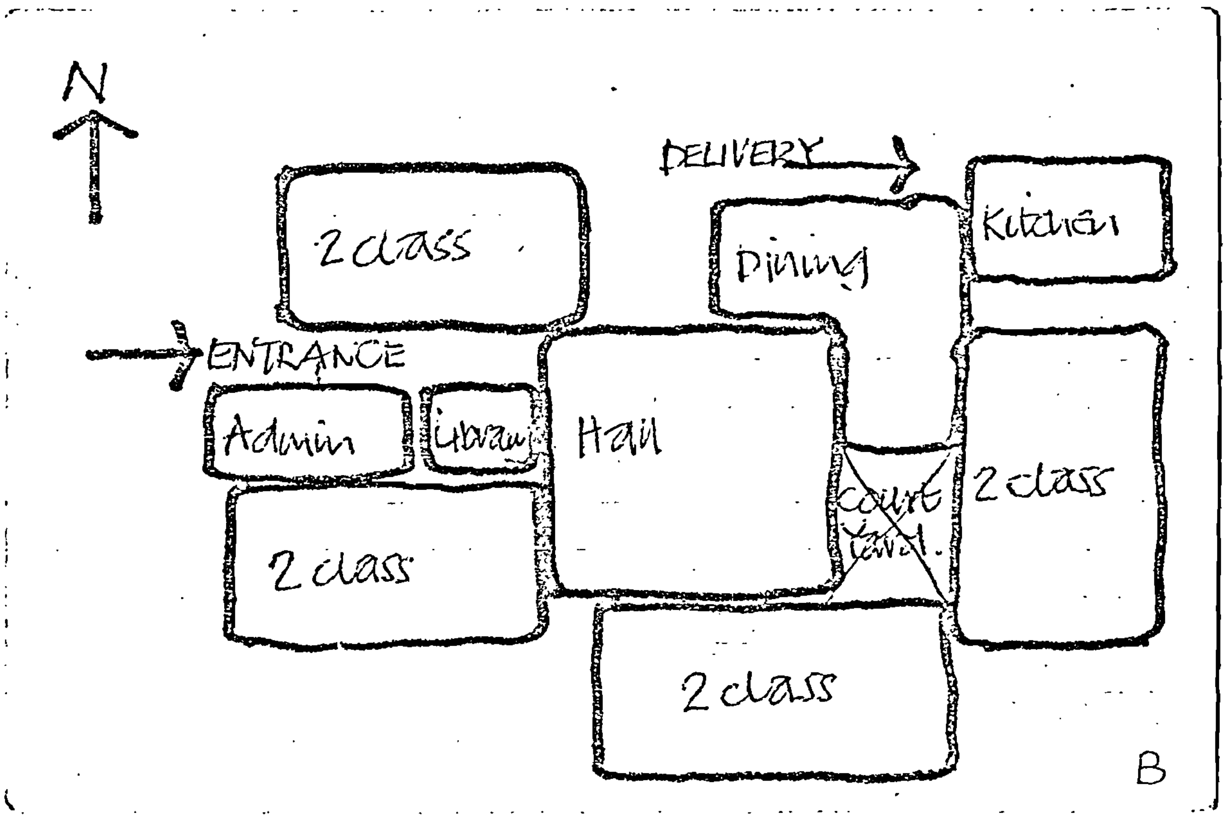
Key + ——— Overall evaluation eigenvector
× — — — Additive utility model index

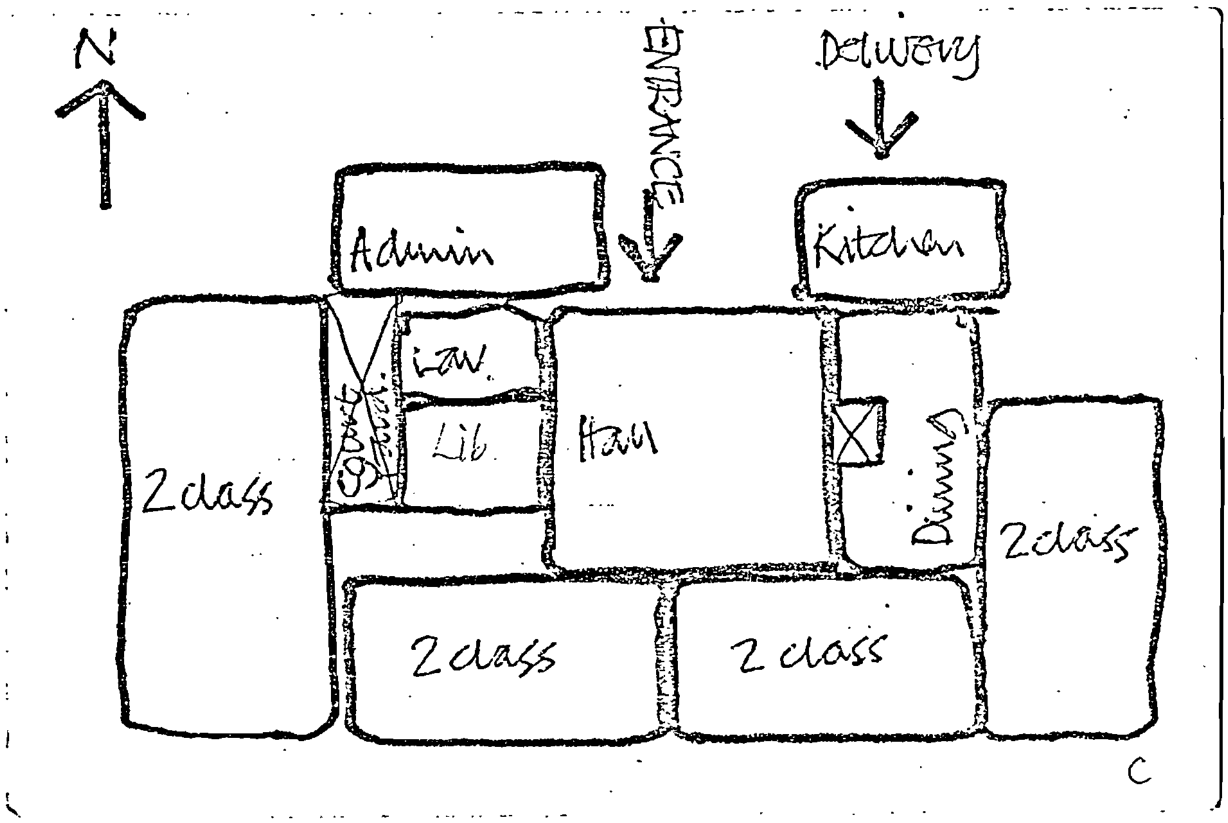
Figure 5.2 The six school plans

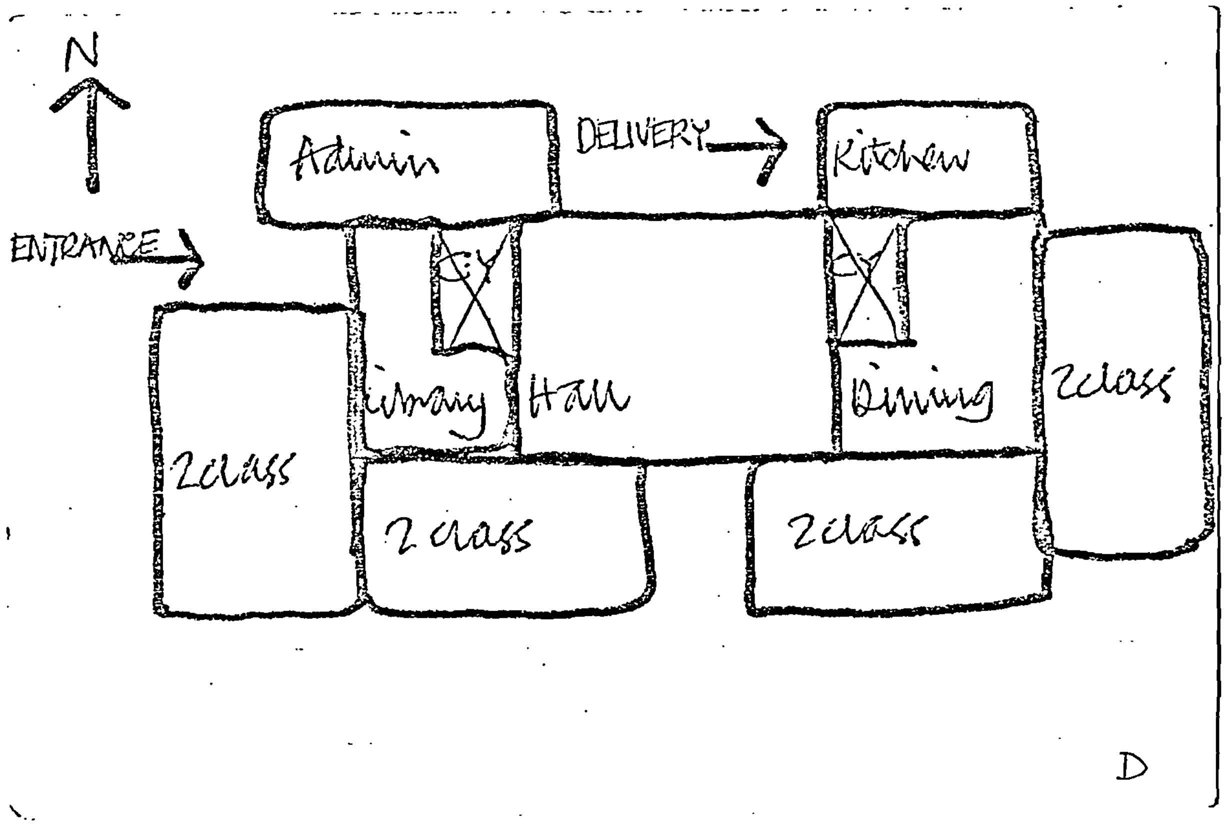
The following six pages show the school plans used in Experiment One.

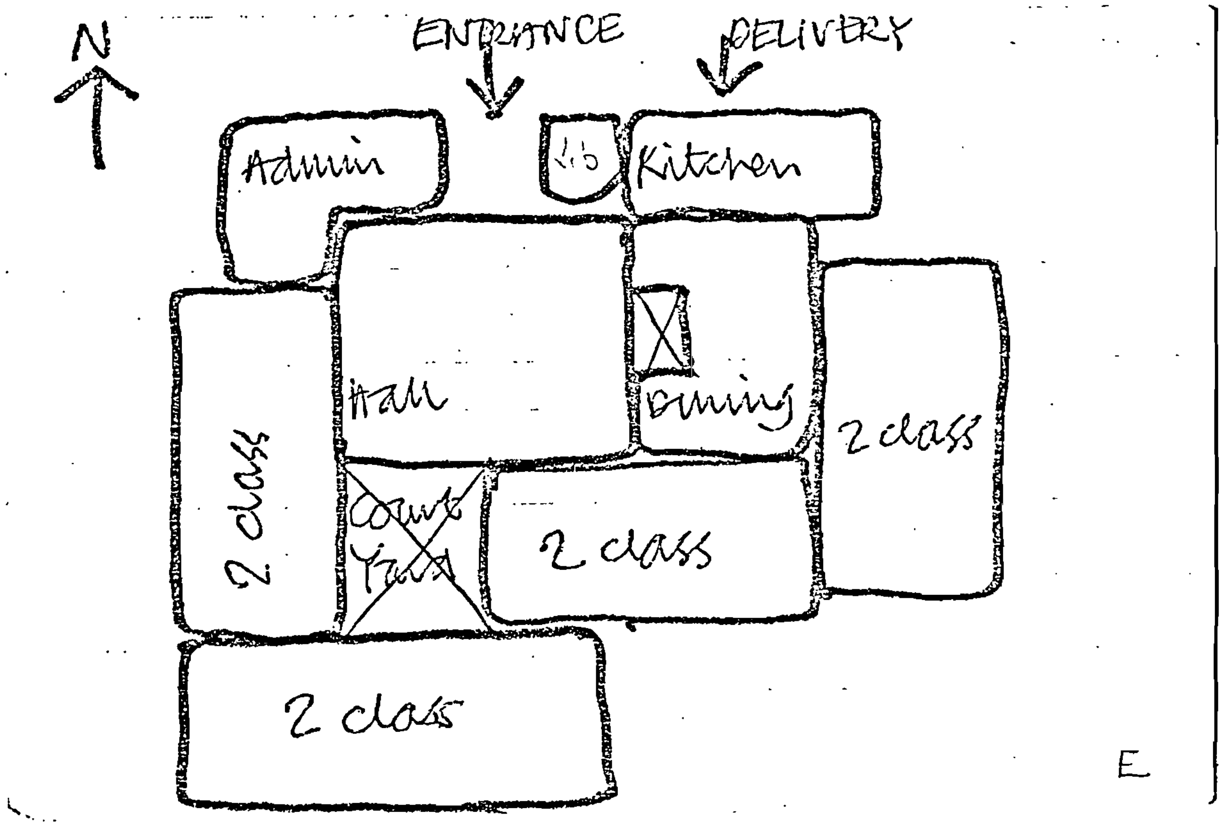


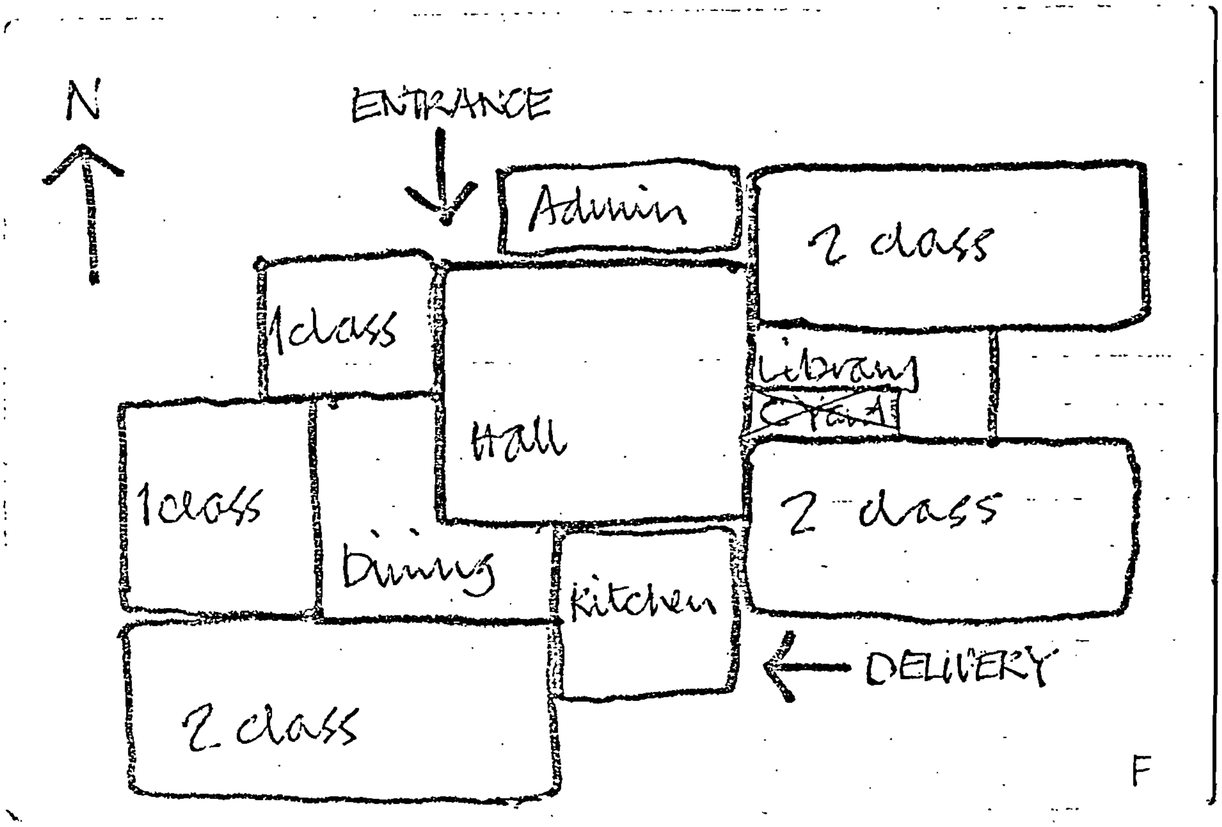
N
2class
Admin
ENTRANCE
Library
Hall
2class
2class
Dining
Kitchen
2class
DELIVERY
A


N
DELIVERY
2 class
Dining
Kitchen
ENTRANCE
Admin
Library
Hall
court yard.
2 class
2 class
2 class
B


N
ENTRANCE
Delivery
Admin
Kitchen
Lav.
Lib
Hall
Dining
2class
2class
2class
2 class
C


N
ENTRANCE
Admin
DELIVERY
Kitchen
Library
Hall
Dining
2class
2class
2 class
2class
2class
D


N
ENTRANCE
DELIVERY
Admin
Lib
Kitchen
Hall
Dining
2 class
2 class
Court Yard
2 class
2 class
E


N
ENTRANCE
Admin
2 class
1class
Hall
Library
1class
2 class
Dining
Kitchen
DELIVERY
2 class
F

CHAPTER 6 EXPERIMENT TWO

6.1 Introduction

Experiment One showed the potential usefulness of Saaty's prioritization and hierarchical cluster analysis in the evaluation of designs, and in the eliciting and rating of attributes. The second experiment makes use of prioritization of attributes in a design context, to elicit designers' priorities. The attempt is made to measure changes in designers' priorities caused by the design process, and to try to find out whether the internal consistency with which attributes are rated is improved as a result of the design process. Additionally the design solutions are evaluated subjectively by the designers themselves, and by independent judges. An attempt is made to discover whether each designer's stated priorities can be observed and measured in his design solution. Hypotheses are also tested about levels of agreement between the evaluations.

For this exercise graduate students of architecture were used as subjects; these were fourth year students at the Liverpool School of Architecture in the first year of the two year BArch course. All had completed a three year undergraduate course and one year of practical training.

A one week design exercise was used as a vehicle for the experiment. The brief (appendix 6.1) was to design a two man Coast Guard Station for mass production. The time-table (appendix 6.2) allowed half a day for background investigation by groups of students, one day for design work, and the remainder for an introduction, a feedback session from the groups' investigations, sessions for the scaling of attributes, assessments of the schemes, and a concluding discussion about the results.

In addition to being an exercise concerned with design method it was also an exercise for learning about glass reinforced polyester (grp) technology. In that context and for the purposes of the experiment, four aspects of the technology were agreed upon jointly by the experimenter and two tutors as being suitable as foci for the background investigations and to serve as attributes for design and evaluation.

The reasons for this were several. First they enabled the year to be divided conveniently into groups for undertaking the background investigations. Second common attributes could be discussed explicitly so that each person would have a common basis for making judgements. Third because all subjects were sharing the same attributes, the paired comparison technique could be performed with the whole year simultaneously; attempting to do it individually would have been excessively time consuming.

6.2 Hypotheses

The experiment was designed to test a number of hypotheses. All are expressed as null hypotheses.

6.2.1 Correlations between the ratings of attributes before and after design

Hypothesis 1 That for each subject there would be no significant difference between his rating of attributes before design compared with his rating after design.

Hypothesis 2 That there would be no significant concordance between the ratings of attributes before design.

Hypothesis 3 That there would be no significant concordance between the ratings of attributes after design.

6.2.2 Internal consistency in prioritization before and after design

Hypothesis 4 That there would be no significant differences between the internal consistency achieved in prioritization of attributes before design compared with after design.

Hypothesis 5 That internal consistency achieved in prioritization of attributes would not improve as a result of performing prioritization, in conjunction with the design process.

Hypothesis 6 That internal consistency achieved in prioritization of attributes would not improve as a result of performing prioritization and being given the results, in conjunction with the design process.

6.2.3 Effect of performing prioritization on design performance

Hypothesis 7 That performing prioritization of attributes would not improve the subjects' design performance.

Hypothesis 8 That performing prioritization of attributes and being given the results would not improve the subjects' design performance.

6.2.4 Consistency and design performance

Hypothesis 9 That there would be no correlation between the total average score achieved by each designer's scheme and the internal consistency measure he achieved in prioritization of attributes.

6.2.5 The relationship between intentions and achievements

Hypothesis 10 That the rating of attributes by each designer would not be correlated with the average scores his design scheme received with respect to each attribute.

6.2.6 Concordance between evaluations of schemes

Hypothesis 11 That there would be no significant concordance within the evaluations of design schemes with respect to each attribute.

Hypothesis 12 That there would be no significant concordance between the evaluations of design schemes with respect to different attributes.

Hypothesis 13 That there would be no significant concordance among the tutors' evaluations.

Hypothesis 14 That there would be no significant correlation between the students' evaluations of the schemes and the tutors' evaluations.

6.3 Experimental Method

In order to test the hypotheses the experiment was designed as follows:

1 Four aspects of the design problem were agreed by tutors and the experimenter as being the foci of the study. The four were:

a <u>structural properties</u> of grp, including such factors as the sizes of the members, anchorage and wind, site conditions and fixing;

b <u>manufacturing requirements</u>, including such factors as the size of the mould, the method of lamination, and the jointing of materials;

c <u>environmental aspects</u>, including such factors as heat loss and condensation;

d <u>interior fixtures and fittings</u>, including such factors as plumbing, wc's and equipment.

2 The subjects were divided into four groups for the purposes of the background investigations. Group A studied aspect a, group B studied aspect b, and so on.

At this stage subjects were not advised of the groupings.

3 The subjects were also divided into three teams X, Y and Z in such a way that each of these teams contained an equal number of members of groups A, B, C and D. Figure 6.1 shows this division in diagramatic form

4 The brief was handed out.

5 Members of teams X and Y made judgements of paired comparisons of the four attributes a, b, c, d described above, using Saaty's prioritization. All four were first displayed on a board. The numerical scale to be used in making judgements was also displayed. The board showing all four attributes was removed and the subjects were shown all possible pairs of attributes in turn. The experimenter kept each pair on display until each subject had written down his weighting. The subjects were shown again the board describing all four aspects and were asked to give a simple ranking of them.

6 Members of team X were given feedback of the results of their scaled paired comparisons. They were shown the normalised eigenvectors and the maximum eigenvalues of their judgements, and the meaning of these was explained to them.

7 The subjects made background investigations into the aspect of grp to which they had been allotted; see section 2 above.

8 Each group in turn gave a presentation of the findings of the background investigation. The presentation comprised at least one A1 sheet of drawings and notes, supported by a verbal description.

9 The subjects then designed their schemes. All worked individually.

10 All schemes were exhibited in the studio. Each scheme was allocated a number, and the number pinned over the subject's name. The schemes were numbered at random from 1 to 26 in order to reduce the ability of the

subjects when evaluating to know whose scheme was which. Although it may be argued that each individual's drawing style would be recognised by his contemporaries it was felt by the experimenter first, that it was likely that not everyone would recognise everyone else's style even if they did recognise several close friends' and two or three distinctive styles; second that even this very modest degree of anonymity would give the evaluators some personal degree of separation from their contemporaries and allow them to rationalise to themselves that they did not know whose scheme they were marking even if they were almost certain that they did know. It is rare, at least at Liverpool, to allow students to mark each other's work, and these very modest precautions were taken because of its novelty. The hope was that because the scheme was associated with just a number and not a person's name, it would be judged simply as submitted, uninfluenced by the designer's past performance or reputation.

11 All the subjects repeated prioritization of the four attributes by scaling paired comparisons as described above.

12 The designers now became evaluators and marked all the schemes. The scale for marking was fully specified in order to increase inter-evaluator reliability and in order to avoid preconceived notions of the narrow band of marking commonly associated with undergraduate assessment. Appendix 6.3 describes the scale. (The use of paired comparisons would of course have been impossible.) Each evaluator marked the schemes for their performance on the particular attribute into which that evaluator as a designer had been responsible for undertaking beckground investigation. Thus members of group A marked the schemes for their performance on attribute a, members of group B on attribute b, and so on.

13 Three independent tutors who had not previously been involved scaled paired comparisons of the four attributes.

14 These three same tutors gave each design scheme an overall mark using the same scale as the student evaluators had used.

15 On the last afternoon a review was held to discuss some of the results. These comprised those which the experimenter was able to calculate at the time.

16 Summary

The experiment may be summarised thus:

Before designing:

i Members of team X scaled paired comparisons of attributes and were given feedback of their results.

ii Members of team Y scaled paired comparisons of attributes but were not given feedback.

iii Members of team Z were not involved in this stage.

After designing, all subjects performed prioritization of attributes.

In evaluation:

i Each subject marked every scheme for its performance with respect to the particular aspect of the design problem (a, b, c or d) to which he had been allocated initially.

ii Three tutors performed paired comparisons of the four attributes a, b, c and d.

iii The three tutors marked the schemes for overall performance.

6.4 Results

6.4.1 Correlations between the ratings of attributes before and after design

Hypothesis 1 was tested by comparing the rank order of attributes computed from each designer's scaled paired comparisons before and after design. This comparison can be made for members of teams X and Y. Table 6.2 shows the data. For significant correlation between two sets of ranks using Spearman's rank correlation coefficient when there are four items being ranked, there must be no differences between the ranks. Of the seven members of team X all but one changed their rating of the four attributes. Of the seven members of team Y all but one changed their rating of the four attributes. The null hypothesis was not entirely rejected but there is evidence to suggest that a high proportion of designers change their priorities as a result of the design process.

Hypothesis 2 was tested by calculating Kendall's coefficient of concordance between the sixteen sets of ranks of attributes before design. Table 6.2 shows the data. The coefficient was found to be 0.329 which was significant ($\alpha = 0.01$). The null hypothesis was rejected.

Hypothesis 3 was tested by calculating Kendall's coefficient of concordance between the twenty sets of ranks of attributes after design. Table 6.2 shows the data. The coefficient was found to be 0.332 which was significant ($\alpha = 0.01$). The null hypothesis was rejected.

6.4.2 Internal consistency in prioritization before and after design

Hypothesis 4 was tested by calculating Mann-Whitney's U between the maximum eigenvalues (internal consistency measures) achieved in scaling paired comparisons of attributes by members of team Z who rated attributes for the first time after designing, with those achieved by members of teams X and Y before design. Table 6.2 shows the data. U was found to be 27 which was not significant. The null hypothesis was not rejected.

Hypothesis 5 was tested by calculating Mann-Whitney's U between the maximum eigenvalues achieved in scaling paired comparisons of attributes by members of team Y before and after design. Table 6.2 shows the data. U was found to be 19.5 which was not significant. The null hypothesis was not rejected.

Hypothesis 6 was tested by calculating Mann-Whitney's U between the maximum eigenvalues achieved in scaling paired comparisons of attributes by members of team X who performed prioritization before and after design. Table 6.2 shows the data. U was found to be 20 which was not significant. The null hypothesis was not rejected.

6.4.3 Effect of performing prioritization on design performance

Hypothesis 7 was tested by calculating Mann-Whitney's U between the total average mark with respect to the four attributes (excluding tutors' marks) by the schemes designed by the members of team Y who performed prioritization before design, and the total average mark received (excluding tutors' marks) by the schemes designed by the members of team Z. Table 6.6 shows the data. U was found to be 21 which was not significant. The null hypothesis was not rejected.

Hypothesis 8 was tested by calculating Mann-Whitney's U between the total average mark with respect to the four attributes received (excluding tutors' marks) by the schemes designed by the members of team X who performed prioritization before design and the total average mark received (excluding tutors' marks) by the schemes designed by the members of team Z. Table 6.6 shows the data. U was found to be 36 which was not significant. The null hypothesis was not rejected.

6.4.4 Consistency and design performance

Hypothesis 9 was tested by calculating Spearman's rank correlation coefficient between the ranks of maximum eigenvectors achieved by each subject in scaling paired comparisons of attributes after design, and the ranks of the total average mark his scheme was given by the assessors (excluding tutors). As there were no significant differences between the maximum eigenvectors before and after design the hypothesis was tested using the results of all twenty subjects who had rated attributes after design. Table 6.5 shows the data. The coefficient was found to be -0.169 which was not significant (one-tailed test). The null hypothesis was not rejected.

6.4.5 The relationship between intentions and achievements

Hypothesis 10 was tested by comparing the rank order of attributes computed from each designer's scaled paired comparisons after design with the rank order of average scores awarded to his scheme with respect to each attribute. Table 6.2 shows the rank order of attributes, table 6.3 shows the marks awarded to design schemes and table 6.4 shows the average mark awarded with respect to each attribute. For significant correlation between two sets of ranks using Spearman's rank correlation coefficient when there are four items being ranked, there must be no differences between the ranks. Of the twenty subjects' results compared, only two, S2.5 and S2.8 showed significant correlations between the two sets of ranks. Two additional alternative ways of testing this hypothesis were also tried. First each assessor's scores were 'normalised'. The mean and standard deviation for each assessor were calculated, and each score expressed using the formula (score - mean)/standard deviation. This was an attempt to cancel the effects of the judges using the specified scale differently from one another. For each attribute the normalised scores of each design scheme were added and then divided by the number of assessors to give an average normalised score. The rank order of attributes computed from each designer's scaled paired comparisons after design was compared with the rank order of average normalised scores awarded to his scheme with respect to each attribute. In only one case (S2.5) out of twenty was there significant correlation between the two sets of ranks. The second additional alternative entailed drawing the profiles of weighting of attributes computed from each designer's scaled paired comparisons, and the profiles of the average scores awarded to each designer's scheme, in the expectation that visual comparison of the two profiles would reveal correlations. Figures 6.2 and 6.3 show these profiles. Perhaps the most conspicuous point to emerge from the profile of average marks awarded to each scheme is the tendency for each scheme to score approximately equal marks with respect to each attribute. This remains true for both good and weak schemes, for example, schemes F and L, and schemes A and G. The only conspicuous exception is scheme U which scored well with respect to attributes a and b, but poorly with respect to attributes c and d. The profile of the relative weighting of attributes on the other hand indicates that the designers did not rate the attributes approximately equal, with the exceptions of S2.11 and S2.17. None of the three means of testing hypothesis 7 gave expected results. The null hypothesis was not rejected.

6.4.6 Concordance between evaluations of schemes

Hypothesis 11 was tested by calculating Kendall's coefficient of concordance between the ranks of marks awarded by assessors with respect to each separate attribute, that is, within each group. Table 6.3 shows the marks awarded tabulated according to attributes (groups). Table 6.7 shows the coefficients of concordance with respect to each attribute (within each group). In all four groups concordance was highly significant. The null hypothesis was rejected.

Hypothesis 12 was tested by calculating Kendall's coefficient of concordance between the ranks of marks awarded by assessors between groups. Table 6.7 shows the coefficients of concordance. In all cases concordance was highly significant. The null hypothesis was rejected.

Hypothesis 13 was tested by calculating Kendall's coefficient of concordance between the ranks of marks awarded by the three tutors. The coefficient was found to be 0.649 which was significant at the 0.01 level. The null hypothesis was rejected.

Hypothesis 14 was tested by calculating Spearman's rank correlation coefficient between the ranks of the total average marks with respect to the four attributes received by each scheme and the ranks of the total marks awarded by the three tutors. Tables 6.3 and 6.4 show the data. The coefficient was found to be 0.689 which was significant at the 0.01 level (one-tailed test). The null hypothesis was rejected.

6.5 Conclusions

The result of testing hypothesis 1 is indicative that a high proportion of subjects who took part in this experiment changed their rating of priorities during the experiment. It seems probable that the change was caused by the design process. However it may be the case that priorities do change over time for no apparent reason, and subsequent experiments 3, 4 and 5 attempt to explore this possibility.

The result of testing hypotheses 2 and 3 were slightly surprising in that it had been expected that designers would differ in the attributes they value, and that these differences would account, in part, for differences in their design schemes. That they account only in part is because the four common attributes were decided by the tutors rather than individually by students. Nevertheless while these attributes do not necessarily represent the main objectives of each designer, the relative weightings of the attributes by the designers give an indication of their relative priorities among given attributes. Furthermore while there is significant agreement between the designers' ranks of the four attributes, there are many differences when the weights attached to the attributes are taken into account, as shown in figure 6.2.

As the designers' priorities may change during the design process, it was expected that before they began designing, a flexible approach to their priorities would manifest itself as inconsistency in scaling paired comparisons of attributes before design, and that these inconsistencies would be resolved during the design process. According to the results of testing hypothesis 4 there is no evidence to support this belief.

The results of testing hypotheses 5 and 6 show that consistency in scaling paired comparisons of attributes was not improved through having performed prioritization in conjunction with the design process, with or without being told the weights of attributes and the degree of internal consistency.

The results of testing hypotheses 7 and 8 show that the performance of the designer, as measured by the marks awarded to his design scheme, was not improved through having performed prioritization in conjunction with the design process, with or without being told the weights of attributes and the degree of internal consistency. Like the previous results (of testing hypotheses 5 and 6) these results are not really surprising. With reference to the work of Abercrombie, discussed in 1.1, it is clear that in her work to encourage decision-makers to become aware of the factors which influence their judgements, she spent a whole term teaching the degree of self-awareness necessary for measurably improved judgements to be made.

While internal consistency in scaling paired comparisons of attributes was generally good (as shown in table 6.2) it had been expected that consistency might be correlated with design ability (as measured by the total average score awarded with respect to each attribute). The result of testing hypothesis 9 shows that this was not the case. A possible explanation is that good designers tend to keep a fairly flexible view of priorities, though as the rank correlation coefficient was near zero there is no indication that design ability is inversely correlated with internal consistency among judgements.

The result of testing hypothesis 10 was a disappointing one. It had been hoped that correlation could be found between the designer's rating of attributes or intentions, and the marks awarded to his scheme with respect to each of the four attributes, that is, his achievements. The first way of testing this hypothesis seemed the most likely to show the expected results but did not in fact do so. Normalising the scores awarded to each scheme attempted to cancel the differences between each assessor's use of the scale for marking. In a sense this is a self defeating procedure because the whole point of specifying the scale fully was to be able to compare the marks awarded with respect to the different attributes. For example, many designers rated attribute 'd' least important of the four. It would be expected therefore that their schemes reflect this low rating and in turn that the marks awarded with respect to attribute 'd' be relatively low. Normalising the scores cancels out any such occurrence. But in any case the expected correlations were not found. Comparing profiles of intentions and achievements (figures 6.2 and 6.3) was another alternative. The tendency has already been noted that each scheme scored approximately equal marks with respect to each attribute. If designers had rated the attributes approximately equal then the lack of correlation between intentions and achievements could be more readily understood; Spearman's rank correlation coefficient would be an unsuitable measure for discriminating between ranks which are approximately equal. However as figure 6.2 shows the weights ascribed to the attributes by each designer varied considerably, with the exceptions of S2.11 and S2.17 who rated all the attributes equal. There was no equivalent variation in the marks achieved by the schemes, and the expected correlations were not found.

The shapes of the profiles of marks awarded, particularly the scoring of approximately equal marks with respect to each attribute, are reflected in the results of testing hypotheses 11 and 12. They confirm that the assessors within each group were in agreement about the relative merits of the schemes with respect to each separate attribute. They also confirm that when the ranks of marks awarded by two or more groups were combined there was still highly significant concordance. Thus schemes which scored well with respect to one attribute also scored well with respect to the other three.

The results of testing hypotheses 13 and 14 come as no surprise in view of the above results. As there was agreement about the relative merit of each scheme between the student evaluators, even when they were marking with respect to different attributes, it would be expected that there be concordance among the three independent assessors. It would also be expected that there be correlation between the students' assessments and the tutors' assessments. Both proved to be true.

The results of testing hypotheses 12, 13 and 14 were not entirely in accordance with the tentative theory developed in chapter 2. The theory suggested that designers would differ in the attributes they value, that these differences would manifest themselves in the design schemes, and that in evaluation different schemes would score well with respect to the attributes most valued by the designer, and not so well with respect to less highly valued attributes. These expectations have not been borne out in this experiment. A probable explanation may be postulated in terms of the existence of a confounding variable, which was not controlled in the evaluation phase of the experiment. No control was exercised over the quantity of information which the drawings conveyed. Cakin (1976) has shown how information additional to basic plans, sections and elevations may cause convergence among judgements. If the design schemes varied in the amount of information they conveyed this would probably account for the recorded level of concordance among the judgements. Neither was control exercised over drawing style. It appears essential in future experiments to control these variables.

Finally a note must be made about the organisation of the experiment. Owing to factors beyond the experimenter's control, not everyone was present for all the sessions. For example, although there were twenty-six schemes submitted there were only nineteen student assessors. Groups contained different numbers of students respectively. Teams contained different numbers of students respectively too. A certain amount of juggling with the figures has therefore been necessary. The tables show this juggling explicitly and it can only be hoped that the results have not been impaired by the lack of control. The converse argument is that excessive control over student attendance might have caused the results to suffer. There is a sense in which this one week exercise, although the subject of a controlled experiment, has been conducted in the way that most design exercises are conducted.

Table 6.1 Allocation of subjects into Teams and Groups

Subject	Team	Group
S2.1	Z	A
S2.2	Y	B
S2.3	X	A
S2.4	X	B
S2.5	X	A
S2.6	Y	*
S2.7	X	D
S2.8	X	C
S2.9	X	B
S2.10	Z	D
S2.11	Y	B
S2.12	Z	B
S2.13	Y	A
S2.14	X	C
S2.15	Z	D
S2.16	Y	D
S2.17	X	D
S2.18	Y	C
S2.19	Z	D
S2.20	X	C
S2.21	Z	A
S2.22	Z	D
S2.23	Z	A
S2.24	Y	D
S2.25	Y	C
S2.26	Z	B

* S2.6 was not allocated to a group

Table 6.2 Ranks of attributes derived from scaled paired comparisons before and after design, and associated consistency measures

Subject	Before design					After design				
	Ranks				λmax	Ranks				λmax
	a	b	c	d		a	b	c	d	
S2.1	-	-	-	-	-	-	-	-	-	-
S2.2	4	1	2	3	8.156	2	1	4	3	4.948
S2.3	3	2	1	4	5.733	1	2	3	4	4.319
S2.4	2	1	3	4	4.119	1.5	1.5	4	3	4.155
S2.5	1	3	2	4	4.638	1	2	3	4	4.621
S2.6	4	2	1	3	4.190	-	-	-	-	-
S2.7	4	1.5	1.5	3	4.060	4	1.5	1.5	3	4.006
S2.8	-	-	-	-	-	1	3	2	4	4.033
S2.9	3	1	2	4	4.479	2	1	3	4	4.735
S2.10	-	-	-	-	-	3	2	1	4	4.321
S2.11	2.5	2.5	2.5	2.5	4.000	2.5	2.5	2.5	2.5	4.000
S2.12	-	-	-	-	-	1	3	2	4	8.568
S2.13	2	1	3	4	4.517	1	2	3	4	4.637
S2.14	1.5	1.5	3	4	4.154	1	2	3	4	4.347
S2.15	-	-	-	-	-	4	1.5	1.5	3	4.044
S2.16	2	1	4	3	4.226	3	1	2	4	4.084
S2.17	2	3	1	4	4.413	2.5	2.5	2.5	2.5	4.000
S2.18	1	3	4	2	6.136	1	4	3	2	4.359
S2.19	-	-	-	-	-	1	3	2	4	6.432
S2.20	3	1	2	4	5.016	-	-	-	-	-
S2.21	-	-	-	-	-	3	1	2	4	4.532
S2.22	-	-	-	-	-	-	-	-	-	-
S2.23	-	-	-	-	-	-	-	-	-	-
S2.24	3	2	1	4	4.664	2	1	3	4	4.742
S2.25	1	2	3	4	4.228	1	3	2	4	4.097
S2.26	-	-	-	-	-	-	-	-	-	-

Table 6.3 Marks awarded by evaluators, tabulated according to attributes with respect to which marks were awarded

Design schemes	Evaluators																					
	attribute a			attribute b					attribute c				attribute d							Tutors		
	S2.3	S2.5	S2.13	S2.2	S2.4	S2.9	S2.11	S2.12	S2.8	S2.14	S2.18	S2.25	S2.7	S2.10	S2.15	S2.16	S2.17	S2.19	S2.24	T1	T2	T3
A	1	12	4	5	9	5	7	6	6	1	7	6	10	5	5	2	3	15	4	6	4	4
B	17	14	16	17	13	9	14	10	8	4	16	10	8	12	8	9	12	15	12	16	9	12
C	16	4	8	15	16	3	9	7	8	1	10	8	10	9	8	6	10	10	15	6	6	4
D	4	4	9	3	9	8	16	13	7	3	10	11	15	12	15	5	9	10	12	11	13	7
E	10	16	9	8	6	15	17	8	8	4	12	11	10	9	4	5	4	10	4	7	6	8
F	15	14	15	11	18	15	18	7	12	8	17	15	15	12	13	16	16	10	15	20	10	11
G	11	1	4	7	4	3	12	8	6	4	6	8	12	5	1	5	2	3	4	3	4	4
H	14	16	11	11	9	11	13	13	16	3	18	9	12	10	2	14	3	15	3	8	8	3
I	12	11	14	10	16	12	16	11	12	13	17	13	15	17	2	14	4	15	14	8	13	15
J	17	11	14	5	8	11	14	10	10	7	9	8	12	10	13	8	8	12	15	9	7	6
K	18	19	15	17	17	18	15	14	12	10	11	7	15	12	14	14	15	15	17	17	5	16
L	18	9	12	13	16	9	12	12	12	6	17	14	14	15	12	9	9	15	16	8	7	4
M	13	7	12	15	16	10	12	11	7	9	7	5	10	10	8	6	15	3	15	7	6	7
N	15	14	16	7	11	14	14	8	12	9	9	9	6	12	14	8	15	5	13	8	4	8
O	8	10	9	5	10	10	12	11	9	3	12	12	6	9	12	8	5	5	15	9	6	7
P	11	15	11	17	10	15	18	14	12	13	16	14	18	14	10	14	16	15	5	20	14	18
Q	16	11	16	14	16	11	14	14	13	8	13	10	15	10	16	13	16	18	16	14	15	16
R	13	7	14	11	8	14	10	11	9	7	10	12	14	7	6	9	5	2	6	7	5	12
S	12	14	10	12	12	7	16	12	12	4	10	8	15	12	8	10	13	15	14	12	3	6
T	8	5	11	9	8	7	7	10	6	4	11	12	6	9	7	13	2	5	13	7	5	4
U	7	12	10	15	9	9	13	12	6	1	7	5	5	7	8	2	5	1	11	9	7	3
V	16	12	15	11	13	10	14	9	8	3	12	6	9	10	12	9	14	2	14	10	10	8
W	9	10	9	1	8	8	15	9	8	6	14	10	10	9	11	10	10	3	7	9	3	10
X	2	18	16	16	16	15	19	11	10	2	13	5	12	10	8	10	15	10	6	17	12	16
Y	6	6	8	3	8	3	5	6	8	6	9	12	6	1	13	6	12	0	12	9	6	4
Z	12	6	11	10	6	9	11	9	9	3	8	11	10	10	14	8	10	1	15	5	8	6

Design schemes are represented by rows; columns show marks awarded by evaluators, tabulated according to groups.

Table 6.4 Average mark awarded to each scheme evaluated with respect to each attribute, and total of average marks

Scheme	Average mark				Rank				Total	
	a	b	c	d	a	b	c	d	mark	Rank
A	5.67	6.40	5.00	6.29	3	1	4	2	23.36	25
B	15.67	12.60	9.50	10.86	1	2	4	3	48.63	8
C	9.33	10.00	6.75	9.71	3	1	4	2	35.79	19
D	5.67	9.80	7.75	11.14	4	2	3	1	34.36	21
E	11.67	10.80	8.75	6.57	1	2	3	4	37.79	16
F	14.67	13.80	13.00	13.86	1	3	4	2	55.33	2
G	5.33	6.80	6.00	4.57	3	1	2	4	22.70	26
H	13.67	11.40	11.50	8.43	1	3	2	4	45.00	11
I	12.33	13.00	13.75	11.57	3	2	1	4	50.65	6
J	14.00	9.60	8.50	11.14	1	3	4	2	43.24	12
K	17.33	16.20	10.00	14.57	1	3	4	2	58.10	1
L	13.00	12.40	12.25	12.86	1	3	4	2	50.51	7
M	10.67	12.80	7.00	9.57	2	1	4	3	40.04	14
N	15.00	10.80	9.75	10.43	1	2	4	3	45.98	9
O	9.00	9.60	9.00	8.57	2.5	1	2.5	4	36.17	18
P	12.33	14.80	13.75	13.14	4	1	2	3	54.02	4
Q	14.33	13.80	11.00	14.86	2	3	4	1	53.99	5
R	11.33	10.80	9.50	7.00	1	2	3	4	38.63	15
S	12.00	11.80	8.50	12.43	2	3	4	1	54.73	3
T	8.00	8.20	8.25	7.86	3	2	1	4	32.31	22
U	9.67	11.60	4.75	5.57	2	1	4	3	31.59	23
V	14.33	11.40	7.25	10.00	1	2	4	3	42.98	13
W	9.33	8.20	9.50	8.57	2	4	1	3	35.60	20
X	12.00	15.40	7.50	10.14	2	1	4	3	45.04	10
Y	6.67	5.00	8.75	7.14	3	4	1	2	27.56	24
Z	9.67	9.00	7.75	10.00	2	3	4	1	36.42	17

Table 6.5 Consistency achieved by each subject in rating attributes after design, and total average mark received by his design scheme

Subjects	λmax	Scheme score	Ranks of λmax	Ranks of scores
S2.2	4.948	48.63	18	7
S2.3	4.319	35.79	9	16
S2.4	4.155	34.36	8	17
S2.5	4.621	37.79	14	13
S2.7	4.006	22.70	3	20
S2.8	4.033	45.00	4	10
S2.9	4.735	50.65	16	5
S2.10	4.321	43.24	10	11
S2.11	4.000	58.10	1.5	1
S2.12	8.568	50.51	20	6
S2.13	4.637	40.04	15	12
S2.14	4.347	45.98	11	8
S2.15	4.044	36.17	5	15
S2.16	4.084	54.02	6	3
S2.17	4.000	53.99	1.5	4
S2.18	4.359	36.63	12	14
S2.19	6.432	54.73	19	2
S2.21	4.532	31.59	13	18
S2.24	4.742	45.04	17	9
S2.25	4.097	27.56	7	19

Table 6.6 Total average mark awarded (see table 6.4), tabulated according to teams

Team	Subjects	Marks
X	S2.3	35.79
	S2.4	34.36
	S2.5	37.79
	S2.7	22.70
	S2.9	50.65
	S2.14	45.98
	S2.17	54.02
	S2.20	32.31
Y	S2.2	48.63
	S2.6	55.33
	S2.11	58.10
	S2.13	40.04
	S2.16	54.02
	S2.18	38.63
	S2.24	45.04
	S2.25	27.56
Z	S2.1	23.36
	S2.10	43.24
	S2.12	50.51
	S2.15	36.17
	S2.19	54.73
	S2.21	31.59
	S2.22	42.98
	S2.23	35.60
	S2.26	36.42

Note: for teams X and Y only those who performed prioritization of attributes before design have been included (see text)

Table 6.7 Evaluation of schemes: coefficients of concordance within and between groups

Group(s)	W	Sig. level
A	0.600	0.01
B	0.523	0.001
C	0.569	0.001
D	0.453	0.001
A + B	0.483	0.001
A + C	0.440	0.001
A + D	0.419	0.001
B + C	0.410	0.001
B + D	0.418	0.001
C + D	0.420	0.001
A + B + C	0.404	0.001
A + B + D	0.411	0.001
A + C + D	0.396	0.001
B + C + D	0.390	0.001
A + B + C + D	0.385	0.001
TUTORS	0.649	0.01

Figure 6.1 Division of the subjects into Teams and Groups

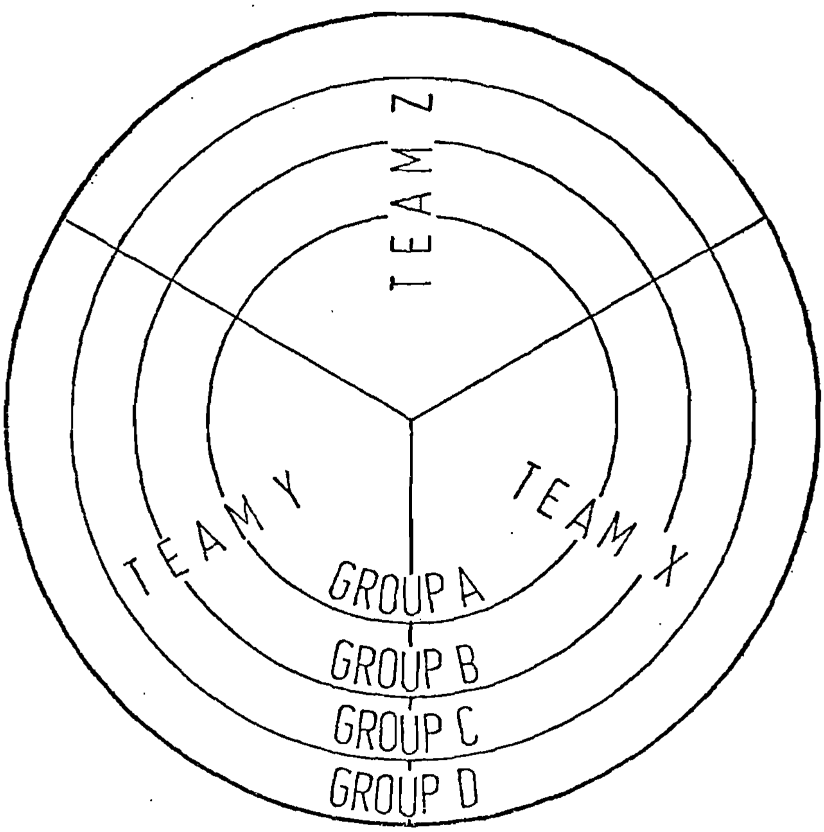


Figure 6.2 Rating of attributes after design

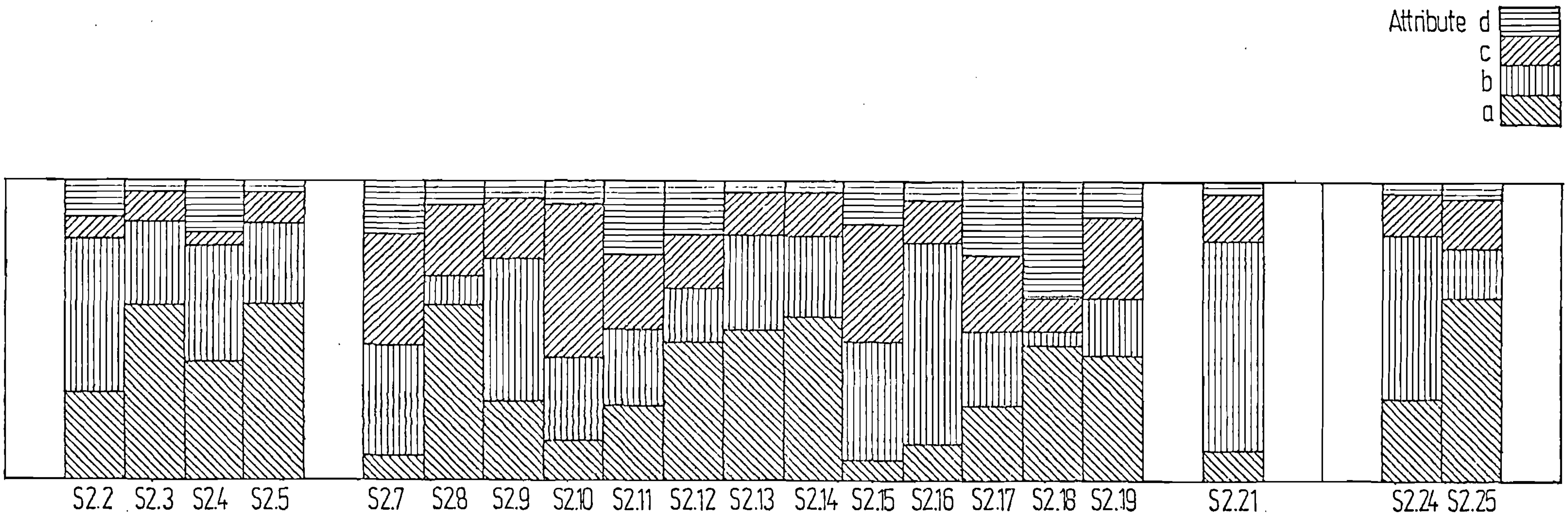
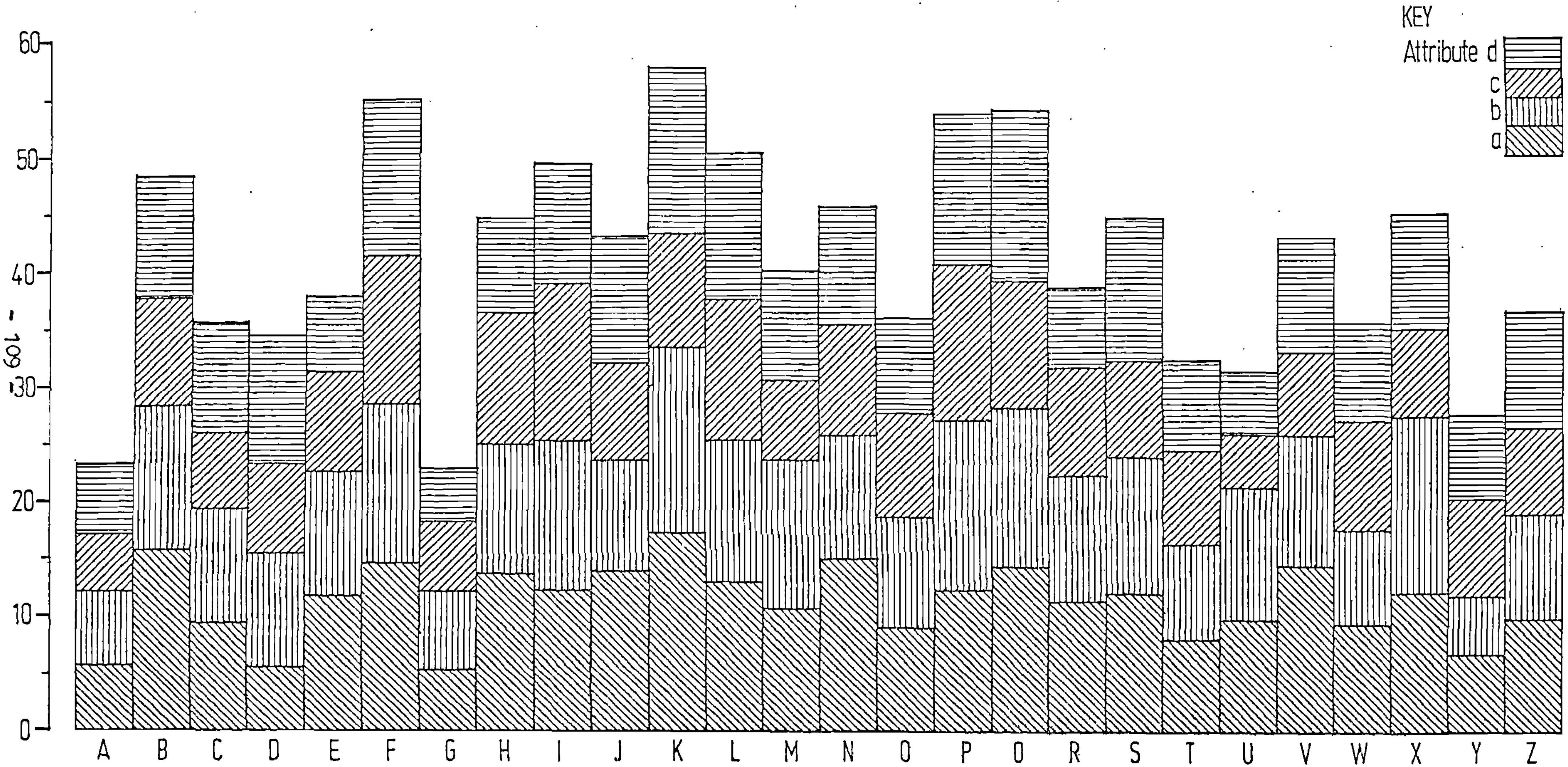


Figure 6.3 Average mark with respect to each attribute received by design schemes



CHAPTER 7 EXPERIMENT THREE

7.1 Introduction

Several important lessons had been learned from the first two exercises about the organisation and running of such experiments, about the means of gathering data, and about the hypotheses it is possible to test. Therefore where the first two experiments have provided this experience, the last three attempt to make use of the findings of the first two; both the formal findings resulting from the testing of hypotheses and the informal findings about organisation and experimental control.

These last three experiments are attempts to explore further designers' values and the evaluation of design. They are all run in a similar, though not identical, fashion. Each comprises six subjects of approximately the same degree of architectural experience. The first uses non-architects, the second students of architecture, the third qualified architects. All three experiments entail the eliciting of dominant attributes from the subjects, an intensive design exercise, the rating of attributes before and after the design process, and subjective evaluations of the design solutions by the subjects themselves. The experimental techniques used are those described in chapter 4. These three experiments are described in this chapter and the two following, respectively.

The lessons to have been learned from the first two experiments may be summarised as follows. A small number of subjects allows a group discussion to take place for the purpose of agreeing the set of the subjects' own dominant attributes. A small number of design schemes allows them to be rated subjectively using paired comparisons, and therefore allows the weighting of attributes and the weighting of designs with respect to each attribute to be combined into an index using an additive utility model. This is a most useful way of exploring subjective evaluation, as shown in the case of subject S1.2, chapter 5. Each subject ought to evaluate each scheme with respect to each attribute, to avoid the difficulty experienced if scales are not used uniformly by each evaluator, as shown in 6.4.5. It is essential that the design schemes are redrawn to cancel both differences in the quantity of information conveyed, and the possible influence of drawing style on the evaluations. Given the importance of these findings Experiment Three attempts to combine the method of evaluation, used successfully in Experiment One, with the eliciting and structuring of designers' intentions studied in Experiment Two. The overall aim is to explore further the theory put forward in chapter 2, and means of testing the theory.

The subjects taking part were six postgraduate research students without formal architectural training. Their backgrounds were in psychology, mining engineering, industrial design, fine art, configurational studies and solar energy. The number of subjects, six, was chosen so that the design schemes produced could be evaluated using the paired comparison technique. Too many design schemes to be compared might have resulted in fatigue for the evaluators (for instance, ten schemes necessitate 45 paired comparisons), while too few might have made the sophistication of scaling paired comparisons superfluous.

A one day intensive design exercise was used as a vehicle for the experiment. The brief (appendix 7.1) was for a two-form entry primary school in Hertfordshire, and was based on the existing brief used in Experiment One. The brief specifies only the areas that are to be provided; it does not specify adjacency conditions or relationships between rooms. A site plan was also provided.

The time-table (appendix 7.2) allowed the first morning for a brain-storming session to find the attributes which the subjects thought important in the planning of the school, for them to agree upon a set of six dominant attributes and for them individually to rate these six in terms of their subjective importance using Saaty's prioritization. In the afternoon each subject designed, individually, a school plan. At the end of the afternoon each subject again scaled paired comparisons of attributes. On the second day the schemes were evaluated formally both with respect to each attribute, and overall, using scaled paired comparisons. Prioritization of attributes was repeated. Triadic comparisons of the school plans were also made.

The number of attributes was chosen as six for three reasons. First is the observation that the human mind is capable of holding only a limited number of concepts at the same time, and this number is 7 ± 2 (Miller, 1956). Second in Experiment One the mean number of attributes used by subjects was $6\frac{2}{3}$. Third although the first two points suggest seven as the most obvious choice, six attributes necessitate only fifteen ($^6C_2$) paired comparisons, as opposed to twenty-one ($^7C_2$) with seven attributes, making the task of scaling rather less demanding.

The use of brainstorming to generate attributes was chosen because the group did not comprise subjects with experience in design. The session gave all subjects a chance to hear others' attributes, which in turn would help them to think of their own, and to propose them without fear of contradiction, since this is explicitly disallowed during brainstorming (Osborn, 1957).

The discussion and eventual agreement on six major attributes to be rated before and after design and to be those with respect to which the schemes were evaluated, was particularly important. Having all six subjects use common attributes enabled the priorities expressed by each designer to be compared with every other designer. Furthermore in the evaluation phase it allowed all their evaluations to be compared directly for concordance. Triadic comparisons of school plans were also made by each evaluator, both to enable hierarchical cluster analysis of the similarity judgements of plans, and to provide a check on whether the originally agreed attributes were actually being used to differentiate between plans.

7.2 Hypotheses

The experiment was designed to test a number of hypotheses. All are expressed as null hypotheses.

7.2.1 Correlations between the ratings of attributes by each subject before design, after design and in evaluation

Hypothesis 1 That for each subject there would be no significant correlation between the rating of attributes before design and the rating after design.

Hypothesis 2 That for each subject there would be no significant correlation between the rating of attributes before design and the rating during evaluation.

Hypothesis 3 That for each subject there would be no significant correlation between the rating of attributes after design and the rating during evaluation.

7.2.2 Concordance between the subjects' ratings of attributes

Hypothesis 4 That there would be no significant concordance between the subjects' ratings of attributes before design.

Hypothesis 5 That there would be no significant concordance between the subjects' ratings of attributes after design.

Hypothesis 6 That there would be no significant concordance between the subjects' ratings of attributes during evaluation.

7.2.3 Differences between the consistency achieved in scaling attributes before design, after design and during evaluation

Hypothesis 7 That there would be no significant difference between the set of consistency measures (eigenvalues) achieved in scaling attributes before design and that achieved after design.

Hypothesis 8 That there would be no significant difference between the set of consistency measures achieved in scaling attributes before design and that achieved during evaluation.

Hypothesis 9 That there would be no significant difference between the set of consistency measures achieved in scaling attributes after design and that achieved during evaluation.

7.2.4 Differences between the consistency achieved in scaling attributes and that achieved in scaling school plans

Hypothesis 10 That there would be no significant difference between the set of consistency measures achieved in scaling attributes before design and that achieved in scaling school plans.

Hypothesis 11 That there would be no significant difference between the set of consistency measures achieved in scaling attributes after design and that achieved in scaling school plans.

Hypothesis 12 That there would be no significant difference between the set of consistency measures achieved in scaling attributes during evaluation and that achieved in scaling school plans.

7.2.5 Correlation between the ranks of consistency measures achieved in scaling attributes before design, after design and during evaluation

Hypothesis 13 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes before design compared with those achieved after design.

Hypothesis 14 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes before design compared with those achieved during evaluation.

Hypothesis 15 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes after design compared with those achieved during evaluation.

7.2.6 Correlation between the ranks of consistency measures achieved in scaling attributes and those achieved in scaling school plans

Hypothesis 16 That there would be no correlation between the ranks of consistency measures achieved in scaling attributes before design and those achieved in scaling school plans.

Hypothesis 17 That there would be no correlation between the ranks of consistency measures achieved in scaling attributes after design and those achieved in scaling school plans.

Hypothesis 18 That there would be no correlation between the ranks of consistency measures achieved in scaling attributes during evaluation and those achieved scaling school plans.

7.2.7 Concordance between evaluations of school plans

Hypothesis 19 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute a.

Hypothesis 20 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute b.

Hypothesis 21 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute c.

Hypothesis 22 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute d.

Hypothesis 23 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute e.

Hypothesis 24 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute f.

Hypothesis 25 That there would be no significant concordance between the subjects' overall ratings of school plans.

Hypothesis 26 That there would be no significant concordance between the ratings of school plans derived from the additive utility model indices.

7.2.8 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

Hypothesis 27 That for each subject there would be no significant correlation between the overall rating of school plans and the rating given by additive utility model indices.

Hypothesis 28 That for each subject there would be no correspondence between the overall rating of school plans and the hierarchical cluster analysis of similarity judgements of school plans.

Hypothesis 29 That for each subject there would be no correspondence between the additive utility model indices and the hierarchical cluster analysis of similarity judgements of school plans.

7.3 Experimental Method

In order to test the hypotheses the experiment was organised in a research studio as follows:

1 The subjects (S3.1, S3.2, S3.3, S3.4, S3.5, S3.6) were given a brief verbal description of how the exercise would be conducted.

2 The subjects were given copies of the site plan and the brief for a two-form entry primary school. The following statement was read to the group:

> "Consider the implications of planning a two-form entry primary school on the given site and to satisfy the given brief. What important attributes or qualities would you take into account in planning the school?"

They were asked as a group to 'brainstorm' (Osborn, 1957) to offer spontaneously the attributes they considered important. Two points were emphasised: first that criticism of the attributes at this stage was ruled out, second that it was the planning that was important as opposed to, for instance, fittings and finishes. At this session 35 attributes were produced (appendix 7.3) in twenty minutes. All attributes were recorded on a blackboard, fully visible to the group.

3 The subjects were asked to combine or to discard attributes in order to end up with the six which they as a group felt were the most representative or expressive of the important attributes in planning the school.

Through group discussion a list of six common attributes was agreed. (An attempt was made to select the six by a voting procedure based on Delphi in which each subject cast votes according to the six attributes he favoured. The total number of votes received by each attribute were then exhibited. Following a brief discussion a second round of voting was held. After the second round it became clear that because of overlaps among attributes such a procedure was unsuited to the task, and an open discussion was used to agree on six major attributes.) The attributes were labelled 'a' to 'f'.

4 Each subject was then asked to work individually in the rating of these six common attributes. Each was given six blank cards, and was asked to copy the attributes onto the cards, one attribute per card. Each subject was also given the numerical scale to be used (appendix 4.1) and a form with instructions on how to work through the cards and to scale all possible pairs of attributes (appendix 4.2). The completed forms were then retained by the experimenter.

5 Each subject proceeded to design a school plan on the given site and to satisfy the given brief. Each worked individually, having been asked specifically not to confer.

6 Having designed a school plan, each subject again rated the six attributes using the same technique as he had before design.

7 The six school plans produced were re-drawn by the experimenter to a standard scale (1:500), orientation, and format (figure 7.2).

8 The redrawn plans were evaluated by the subjects for their overall merit as schools. Each subject worked individually. Each was given the set of school plans, the numerical scale to be used (appendix 4.1) and a form with instructions on how to work through and to scale all possible pairs of plans (appendix 4.4).

9 The redrawn plans were then evaluated by the subjects with respect to each of the six attributes. Each subject worked individually. Each was given the six redrawn plans, the numerical scale to be used (appendix 4.1) and six forms with instructions on how to work through and to scale all possible pairs of plans (appendix 4.3). Each form stated at the top the name of the attribute with respect to which the plans were to be scaled. The six forms were completed each in turn, according to the alphabetical order of the six attributes.

10 Paired comparisons of the six attributes were scaled for a third time using the same technique as before.

11 Finally triadic comparisons of the school plans were made. Each subject worked individually, using a form which gave instructions on how to work through all possible triads of plans (appendix 4.5). In differentiating between plans subjects were asked to state the attributes which they had used; they were not constrained that these attributes necessarily be chosen from the six common attributes.

7.4 Results

7.4.1 Correlations between the ratings of attributes by each subject before design, after design and in evaluation

Hypotheses 1, 2 and 3 were tested by calculating Spearman's rank correlation coefficient between the sets of ranks of attributes given by each subject before design, after design and during evaluation. Table 7.2 shows the data. The following tabulation shows for each hypothesis (H) and each subject the pairs of sets of ranks being compared, the rank correlation coefficient and the significance level (one-tailed test).

Subject	H	Before design	After design	During eval.	Correlation coefficient	Sig. level
S3.1	1 2 3	x x 	x x	 x x	0.486 0.471 0.829	NS NS 0.05
S3.2	1 2 3	x x 	x x	 x x	-0.657 -0.486 0.943	NS NS 0.01
S3.3	1 2 3	x x 	x x	 x x	0.643 0.414 0.886	NS NS 0.05
S3.4	1 2 3	x x 	x x	 x x	1.000 0.943 0.943	0.01 0.01 0.01
S3.5	1 2 3	x x 	x x	 x x	1.000 1.000 1.000	0.01 0.01 0.01
S3.6	1 2 3	x x 	x x	 x x	0.771 0.600 0.943	NS NS 0.01

Null hypothesis 1 was not rejected.

Null hypothesis 2 was not rejected.

Null hypothesis 3 was rejected.

7.4.2 Concordance between the subjects' ratings of attributes

Hypothesis 4 was tested by calculating Kendall's coefficient of concordance between the six sets of ranks of attributes before design. Table 7.2 shows the data. The coefficient was found to be 0.502 which was significant ($\alpha = 0.01$). The null hypothesis was rejected.

Hypothesis 5 was tested by calculating Kendall's coefficient of concordance between the six sets of ranks of attributes after design. Table 7.2 shows the data. The coefficient was found to be 0.310 which was not significant. The null hypothesis was not rejected.

Hypothesis 6 was tested by calculating Kendall's coefficient of concordance between the six sets of ranks of attributes during evaluation. Table 7.2 shows the data. The coefficient was found to be 0.321 which was not significant. The null hypothesis was not rejected.

7.4.3 Differences between the consistency achieved in scaling attributes before design, after design and during evaluation

Hypotheses 7, 8 and 9 were tested by calculating Mann-Whitney's U between the pairs of sets of consistency measures under consideration. Table 7.3 shows the data. The following tabulation shows for each hypothesis (H) the pairs of sets of eigenvalues being compared, the corresponding value of U and the probability under the null hypothesis.

H	Attributes			U	Probability under Ho
	Before design	After design	During eval.		
7	x	x		10	0.120
8	x		x	12	0.197
9		x	x	14	0.294

Null hypothesis 7 was not rejected.
Null hypothesis 8 was not rejected.
Null hypothesis 9 was not rejected.

7.4.4 Differences between the consistency achieved in scaling attributes and that achieved in scaling school plans

Hypotheses 10, 11 and 12 were tested by calculating Mann-Whitney's U between the pairs of sets of consistency measures under consideration. Table 7.3 shows the data. The following tabulation shows for each hypothesis (H) the pairs of sets of eigenvalues being compared, the corresponding value of U and the probability under the null hypothesis.

H	Attributes			Plans	U	Probability under Ho
	Before design	After design	During eval.			
10	x			x	11	0.155
11		x		x	17	0.469
12			x	x	14	0.294

Null hypothesis 10 was not rejected.
Null hypothesis 11 was not rejected.
Null hypothesis 12 was not rejected.

7.4.5 Correlation between the ranks of consistency measures achieved in scaling attributes before design, after design and during evaluation

Hypotheses 13, 14 and 15 were tested by calculating Spearman's rank correlation coefficient between pairs of sets of ranks of consistency measures. Table 7.3 shows the data. The following tabulation shows for each hypothesis (H) the pairs of sets of ranks being compared, the rank correlation coefficient and the significance level (one-tailed test).

H	Attributes			Correlation coefficient	Sig. level
	Before design	After design	During eval.		
13	x	x		0.429	NS
14	x		x	0.743	NS
15		x	x	0.771	NS

Null hypothesis 13 was not rejected.
Null hypothesis 14 was not rejected.
Null hypothesis 15 was not rejected.

7.4.6 Correlation between the ranks of consistency measures achieved in scaling attributes and those achieved in scaling school plans

Hypotheses 16, 17 and 18 were tested by calculating Spearman's rank correlation coefficient between pairs of sets of consistency measures. Table 7.3 shows the data. The following tabulation shows for each hypothesis (H) the pairs of sets of ranks being compared, the rank correlation coefficient and the significance level (one-tailed test).

H	Attributes			Plans	Correlation coefficient	Sig. level
	Before design	After design	During eval.			
16	x			x	0.543	NS
17		x		x	0.771	NS
18			x	x	0.886	.05

Null hypothesis 16 was not rejected.
Null hypothesis 17 was not rejected.
Null hypothesis 18 was rejected.

7.4.7 Concordance between evaluations of school plans

Hypotheses 19 to 24 were tested by calculating Kendall's coefficient of concordance (W) between the sets of ranks under consideration. Table 7.4 shows the data. The following tabulation shows for each hypothesis (H) the sets of ranks being compared, the corresponding value of W and the significance level.

H	Evaluation of plans with respect to attributes						W	Sig. level
	a	b	c	d	e	f		
19	x						0.237	NS
20		x					0.546	.01
21			x				0.216	NS
22				x			0.198	NS
23					x		0.329	NS
24						x	0.501	.01

Null hypothesis 19 was not rejected.
Null hypothesis 20 was rejected.
Null hypothesis 21 was not rejected.
Null hypothesis 22 was not rejected.
Null hypothesis 23 was not rejected.
Null hypothesis 24 was rejected.

Hypothesis 25 was tested by calculating Kendall's coefficient of concordance between the sets of ranks of the subjects' overall ratings of school plans. The coefficient was found to be 0.067 which was not significant. The null hypothesis was not rejected.

Hypothesis 26 was tested by calculating Kendall's coefficient of concordance between the sets of ranks of the subjects' ratings of school plans given by the additive utility model indices. The coefficient was found to be 0.181 which was not significant. The null hypothesis was not rejected.

7.4.8 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

Hypothesis 27 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of overall ratings of school plans and the ranks of additive utility model indices. Tables 7.5 and 7.6 show the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S3.1	0.657	NS
S3.2	0.143	NS
S3.3	0.771	NS
S3.4	0.314	NS
S3.5	1.000	.01
S3.6	0.200	NS

The null hypothesis was not rejected, except in the case of S3.5 where there was significant correlation.

Hypotheses 28 and 29 were tested by comparing for each subject the hierarchical cluster analysis of similarity judgements with both the overall ratings of school plans, and the additive utility model indices. Figure 7.1 shows the data. The following tabulation gives a verbal estimate on a four point scale, 'high', 'medium', 'low', 'no', of the correspondence for each subject between the hierarchical cluster analysis, the overall rating and the indices.

Subject	Correspondence with hierarchical cluster analysis	
	Overall rating	Index
S3.1	Medium correspondence. Cluster A-D: both rated highly. Cluster B-E: both rated low. Plan C rated nearer B-E.	Low correspondence. Cluster B-E: both rated low.
S3.2	Low correspondence. Cluster B-E: both rated highly.	Low correspondence. Cluster C-F: both rated low. Plan D rated close to C-F.

	Overall rating	Index
S3.3	Medium correspondence. Cluster C-F: both rated highly. Cluster A-B: both rated low.	No correspondence.
S3.4	High correspondence. Cluster A-D-E: all rated highly. Cluster B-F: both rated low.	No correspondence.
S3.5	Medium correspondence. Cluster B-E: both rated highly. Cluster A-D: both rated low.	Medium Correspondence. Cluster B-E: both rated highly. Cluster A-D: both rated low.
S3.6	Medium correspondence. Cluster C-F: both rated highly. Cluster B-E-A: all rated low.	High correspondence. Cluster C-F-D: all rated highly. Cluster B-E-A: all rated low.

Null hypothesis 28 was rejected.

Null hypothesis 29 was not entirely rejected.

7.5 Conclusions

7.5.1 The effect of the design process on rating attributes

The testing of hypotheses 1, 2 and 3 provided some extremely interesting results. The results imply that there are two kinds of subjects. S3.4 and S3.5 both exhibited a significant degree of correlation between all three sets of ratings of attributes: before design, after design and during evaluation. They had a relatively fixed set of priorities, which did not change, either as a result of the design process, or through the evaluation of alternative designs. The other four subjects, S3.1, S3.2, S3.3 and S3.6, did not exhibit a significant degree of correlation between the ratings of attributes before and after design; the design process caused measurable changes in their priorities. It is of particular interest that these four subjects (like the other two) did exhibit a significant degree of correlation between the ratings of attributes after design and during evaluation. Their priorities, having changed during the design process, then remained constant.

These are important findings for several reasons. First they demonstrate that judgements of attributes can remain significantly constant during an intensive design exercise of this kind. Second they are indicative that prioritization is a useful technique for measuring the changes in priorities which some designers experience. Third it seems highly probable from the results of testing hypothesis 3 that the performing of prioritization does not of itself affect the rating of priorities.

7.5.2 Concordance between the subjects' ratings of attributes

Before design there was a surprisingly high coefficient of concordance between the subject's ratings of attributes, significant at the 0.01 level. The subjects were in agreement about the relative importance of attributes for a school plan. After design and during evaluation the coefficient of concordance was not significant; the subjects did not share the same priorities. Taken together with the previous inferences (section 7.5.1) it may be seen that there was concordance among priorities between the subjects before design but that the rating of attributes by four of the six subjects changed as a result of the design process leading to lack of significant concordance among the rating of attributes after design. Thus the design process caused in some cases changes in the rating of priorities but these changes were experienced in a different way by each designer. There was divergence in the rating of priorities caused by the design process.

7.5.3 The effect of the design process on consistency

The results of testing hypotheses 7 and 8 were disappointing. It had been expected that before design the subjects would have had a fluid or dynamic opinion of the relative importance of attributes and that this would manifest itself in the form of relatively inconsistent judgements. During the design process manipulation of the alternatives and decision making about the relative importance of attributes might be expected to result in the replacement of the fluid dynamic model by a definite proposal encapsulating the decisions taken. For this reason consistency might have been expected to improve as a result of the design process. This belief however is not borne out by the figures: consistency does not improve. Neither does it improve as a result of evaluating alternatives: the result of testing hypothesis 9 is as expected. An additional conclusion which may be drawn is that successive attempts at performing prioritization do not of themselves lead to improvements in internal consistency.

7.5.4 Consistency in scaling attributes and school plans

The results of testing hypotheses 10, 11 and 12 show that the subjects were no less consistent in scaling school plans for overall merit than they had been in scaling attributes. This is an encouraging result, considering that the subjects were non-architects, and indicates that they were able to scale the school plan drawings with a degree of consistency comparable with that achieved in scaling verbally stated attributes.

7.5.5 The ranks of consistency measures in scaling attributes

The results of testing hypotheses 13, 14 and 15 showed that those who were most consistent in scaling attributes before design did not necessarily remain so after design.

7.5.6 The ranks of consistency measures in scaling attributes and school plans

The results of testing hypotheses 16, 17 and 18 again enable comments to be made about the inherent level of consistency which each subject is able to achieve in scaling paired comparisons. Although null hypothesis 18 was rejected at the 0.05 level of significance this is scant evidence in the context of the results of testing hypotheses 13, 14, 15, 16 and 17 to support the notion that some subjects are inherently more consistent than others in scaling paired comparisons.

7.5.7 Concordance between evaluations of school plans

The results of testing hypotheses 25 and 26 are as expected. It is believed that because the subjects differ in their rating of priorities so when they are evaluating alternative designs their different priorities cause them to prefer different schemes. This belief is strongly supported by the findings of this experiment.

Conversely in the evaluation of alternative designs with respect to individual attributes it had been expected that there would be significant concordance. Such a view is supported by the results of testing hypotheses 20 and 24 where significant concordance was recorded. However with respect to the other four attributes there was not significant concordance between the evaluations. This suggests that the implications of the attributes are not understood in the same way by each evaluator, which may be demonstrated by looking at the way individual subjects rated the school plans. For example, consider the attribute 'd', "Good access to common areas". Evaluators differed considerably in their rating of plans with respect to this attribute: plan D received different ranks from almost all the evaluators. Looking at the plans it may be surmised that some subjects took the attribute to imply that common areas should be planned adjacent to classrooms, others that it implied compactness, the more compact a scheme the more easy it is to get from classrooms to common areas, and others that it implied that the journey should be pleasant or enjoyable or that there should be clear circulation routes. It seems probable that some or all of these different interpretations account for the lack of concordance between the evaluations.

7.5.8 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

The conclusions which may be drawn from the results of testing hypotheses 27, 28 and 29 concern the comparison between three methods of evaluation of the same stimulus items, the school plans. The following tabulation is a combination of the results of testing hypotheses 27, 28 and 29. For each of the three comparisons between the overall merit ratings, the additive utility model indices and the hierarchical cluster analysis, each subject is given a verbal estimate on a four point scale, 'high', 'medium', 'low', 'no', of the degree of correspondence. In this tabulation the rank correlation coefficients found in testing hypothesis 27 are converted to the verbal description thus: correlation significant at 0.01 level, 'high'; correlation significant at 0.05 level, 'medium'; positive correlation, 'low'; negative correlation, 'no'.

	Degree of correspondence between pairs of results		
Overall ratings	x	x	
Indices	x		x
Cluster analysis		x	x
S3.1	low	medium	medium
S3.2	low	low	low
S3.3	low	medium	no
S3.4	low	high	no
S3.5	high	medium	medium
S3.6	low	medium	high

In this tabulation it should be remembered that it takes only one set of results to fail to correspond with either of the other two sets for this to affect two of the three pairs. Because in all cases except S3.2 there is only one set of results which does not correspond to a medium or high degree with the other two, this set can be identified for each subject thus:

S3.1 Utility model indices failed to correspond
S3.3 Utility model indices failed to correspond
S3.4 Utility model indices failed to correspond
S3.6 Overall ratings failed to correspond

S3.2's three sets of results exhibited little correspondence.
S3.5's three sets of results exhibited the most positive correspondence, with high or medium correspondence throughout.

Additionally the tabulation shows that the most positive degrees of correspondence between sets of results occur between the overall ratings and the cluster analysis.

In the case of S3.5 the additive utility model indices do give a precise account of the subject's overall preferences among design alternatives taking into consideration the weighting of attributes, and the weighting of plans with respect to each attribute. In the cases of the other subjects however the results are less encouraging in this respect. The triadic comparisons help to explain why the indices are or are not correlated significantly with the overall ratings and the cluster analysis. Subject S3.5 used attributes 'a', 'b', 'c', 'd' and 'e' to differentiate between plans, and also rated these the most important five of the six attributes. Conversely subject S3.6 did not use any of the original six attributes when making similarity judgements. Subject S3.4 used attributes 'a', 'b', 'd', 'e' and 'f' but did not include attribute 'c' which he had in fact rated the most important. Subject S3.2 did use all six attributes but added 'linearity' as a seventh, which may account for the lack of correspondence. Unfortunately this is not an entirely sufficient explanation however; subject S3.1 used all six attributes in the triadic comparisons but still did not achieve significant correlation between the overall merit rating and the additive utility model indices. In conclusion it seems probable that the agreement of six common attributes, if they are not fully endorsed by each subject, may give rise to discrepancies in the evaluations, but that even for subjects who do apparently endorse them there may be discrepancies. The measure of consistency is a useful guide to the probability of such discrepancies arising, and as table 7.3 shows the consistency achieved by the subjects in scaling attributes during evaluation is quite poor (as comparisons between table 7.3 and tables 8.3 and 9.5 demonstrate). The relative lack of consistency achieved by subjects in this experiment may result in part because not all the subjects did endorse fully the common attributes, and these two factors combined may have led to the discrepancies recorded.

Experiment Five attempts to remedy these problems by eliciting from the subjects their own individually expressed attributes. A minor improvement, which is tried in Experiment Four, is to amend the order in which evaluations are performed, so that overall merit ratings are obtained only after the subjects have scaled the plans with respect to each separate attribute. It is hoped that greater correspondence may thereby be obtained between the indices and the overall merit ratings, and in turn that overall preferences can be explained accurately by the weighting of attributes and weighting of plans with respect to each attribute.

Table 7.1 Description of attributes for school planning

a	Climatic factors: energy consumption
b	Noise (internal)
c	Classroom planning (flexible, open space, supervision)
d	Good access to common areas: hall, dining room, library, administration
e	Easy access to outside
f	Views to outside

Table 7.2 Rating of attributes for a school plan before and after design and during evaluation, and coefficient of concordance between judgements (W).

	Attri-butes	Eigenvectors						Ranks						Concordance	
		S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	W	Sig.level
Before design	a	.321	.409	.133	.151	.066	.491	2	1	4	3	5	1		
	b	.158	.302	.365	.282	.221	.078	3	2	1	2	2	4		
	c	.353	.110	.085	.416	.509	.137	1	3	5	1	1	3	.502	.01
	d	.089	.094	.254	.029	.093	.207	4	4	2	6	3	2		
	e	.054	.040	.142	.082	.083	.023	5	6	3	4	4	6		
	f	.025	.046	.022	.038	.028	.064	6	5	6	5	6	5		
After design	a	.106	.039	.175	.144	.061	.130	4	5	2.5	3	5	3		
	b	.272	.033	.175	.218	.192	.062	2	6	2.5	2	2	4		
	c	.158	.164	.122	.505	.470	.495	3	3	5	1	1	1	.310	NS
	d	.337	.437	.129	.028	.167	.232	1	1	4	6	3	2		
	e	.099	.256	.377	.061	.079	.025	5	2	1	4	4	6		
	f	.027	.070	.023	.044	.029	.057	6	4	6	5	6	5		
During eval.	a	.157	.036	.267	.238	.036	.076	3	5	1	2	5	4		
	b	.354	.029	.169	.132	.228	.161	1.5	6	3	3	2	3		
	c	.073	.258	.156	.462	.492	.452	4	2	4.5	1	1	1	.321	NS
	d	.354	.481	.156	.032	.143	.258	1.5	1	4.5	6	3	2		
	e	.038	.130	.220	.077	.077	.021	5	3	2	4	4	6		
	f	.024	.067	.031	.058	.024	.033	6	4	6	5	6	5		

Table 7.3 Consistency achieved in scaling attributes and school plans

Subject	Consistency in scaling attributes						Consistency in scaling plans	
	Before design		After design		During evaluation			
	λmax	rank	λmax	rank	λmax	rank	λmax	rank
S3.1	6.904	2	6.550	1	6.751	2	6.426	1
S3.2	6.605	1	6.559	2	6.539	1	6.455	2
S3.3	7.867	6	6.656	3	7.200	5	7.089	5
S3.4	7.233	5	7.016	5	6.888	4	6.586	3
S3.5	7.058	4	6.763	4	6.851	3	6.882	4
S3.6	6.980	3	7.315	6	7.876	6	8.342	6

Table 7.4 Rating of school plans with respect to attributes

Attribute	Plan	Subject											
		Eigenvectors						Ranks					
		S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	S3.1	S3.2	S3.3	S3.4	S3.5	S3.6
a	A	.468	.256	.271	.320	.126	.053	1	3	2	2	4.5	6
	B	.048	.280	.113	.050	.126	.074	5	1	4	5	4.5	5
	C	.127	.053	.320	.068	.447	.243	3	5	1	4	1	2
	D	.261	.108	.168	.360	.130	.098	2	4	3	1	3	3
	E	.048	.260	.076	.182	.136	.078	5	2	5	3	2	4
	F	.048	.043	.053	.021	.034	.453	5	6	6	6	6	1
b	A	.064	.090	.105	.034	.062	.057	5	4	4	5	5	4
	B	.176	.163	.098	.073	.249	.033	2	2	5	4	2	6
	C	.144	.053	.372	.167	.152	.312	3	5	1	2	4	2
	D	.025	.023	.039	.029	.029	.153	6	6	6	6	6	3
	E	.119	.152	.127	.129	.278	.034	4	3	3	3	1	5
	F	.473	.518	.259	.568	.231	.411	1	1	2	1	3	1
c	A	.398	.371	.229	.073	.096	.030	1	1	2	5	5	6
	B	.037	.123	.152	.110	.218	.044	6	3	4	3	2	4
	C	.148	.101	.192	.197	.134	.196	3	4	3	2	3	3
	D	.289	.313	.316	.447	.027	.429	2	2	1	1	6	1
	E	.077	.068	.074	.070	.422	.038	4	5	5	6	1	5
	F	.051	.024	.038	.104	.102	.262	5	6	6	4	4	2
d	A	.447	.409	.251	.066	.070	.093	1	1	2	4	4	4
	B	.062	.316	.087	.036	.062	.044	5	2	4.5	6	5	6
	C	.102	.069	.355	.261	.302	.242	3	5	1	2	2	2
	D	.290	.085	.078	.423	.050	.136	2	4	6	1	6	3
	E	.030	.091	.142	.162	.363	.057	6	3	3	3	1	5
	F	.068	.030	.087	.053	.155	.428	4	6	4.5	5	3	1
e	A	.140	.132	.288	.235	.040	.080	3	3	1	3	5	4
	B	.199	.418	.105	.037	.232	.019	2	1	4	5	2	6
	C	.542	.282	.270	.297	.219	.355	1	2	2	1	4	2
	D	.026	.026	.038	.106	.039	.105	5.5	6	6	4	6	3
	E	.067	.039	.079	.035	.245	.059	4	5	5	6	1	5
	F	.026	.103	.219	.291	.225	.382	5.5	4	3	2	3	1
f	A	.225	.227	.251	.169	.109	.134	2	2	2	3	5	3
	B	.035	.082	.213	.035	.276	.021	6	5.5	3	5	1	6
	C	.505	.376	.305	.483	.175	.432	1	1	1	1	3	1
	D	.059	.082	.051	.095	.054	.079	4	5.5	6	4	6	5
	E	.037	.147	.086	.032	.267	.091	5	3	4	6	2	4
	F	.139	.088	.064	.186	.119	.243	3	4	5	2	4	2

Table 7.5 Rating of school plans given by additive utility model indices, corresponding ranks and coefficient of concordance (W)

Plan	Subjects												Concordance	
	Indices						Ranks							
	S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	W	Sig. level
A	.294	.337	.237	.144	.082	.057	1	1	2	4	5	4		
B	.103	.258	.114	.078	.202	.043	5	2	5	6	2	6		
C	.151	.124	.303	.189	.181	.241	4	4	1	2	3	3	.181	NS
D	.176	.135	.123	.323	.036	.266	3	3	4	1	6	2		
E	.069	.090	.096	.103	.353	.047	6	5	6	5	1	5		
F	.207	.056	.128	.163	.147	.345	2	6	3	3	4	1		

Table 7.6 Overall rating of school plans: eigenvectors, corresponding ranks and coefficient of concordance (W)

Plan	Subjects													
	Eigenvectors						Ranks						Concordance	
	S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	S3.1	S3.2	S3.3	S3.4	S3.5	S3.6	W	Sig. level
A	.444	.095	.135	.197	.047	.053	1	4	4	3	5	5		
B	.039	.390	.074	.061	.249	.081	5	1	6	5	2	3		
C	.136	.164	.305	.091	.178	.239	3	3	1	4	3	2	.070	NS
D	.303	.034	.195	.342	.029	.048	2	6	3	1	6	6		
E	.039	.245	.088	.276	.421	.072	5	2	5	2	1	4		
F	.039	.073	.203	.034	.076	.508	5	5	2	6	4	1		

Table 7.7 Consistency achieved in scaling plans with respect to attributes

Subject	Attribute					
	λmax					
	a	b	c	d	e	f
S3.1	6.179	6.739	6.337	6.359	7.011	6.366
S3.2	6.533	6.824	6.406	6.705	6.529	6.052
S3.3	6.510	7.182	7.504	6.453	6.710	8.075
S3.4	7.172	6.660	7.108	6.832	6.230	6.868
S3.5	6.153	6.684	7.151	6.665	6.025	6.750
S3.6	7.197	6.679	6.652	6.503	7.708	7.061

Figure 7.1 Comparison between each subject's similarity judgments of school plans, the weighting given in overall evaluation and the additive utility model indices.

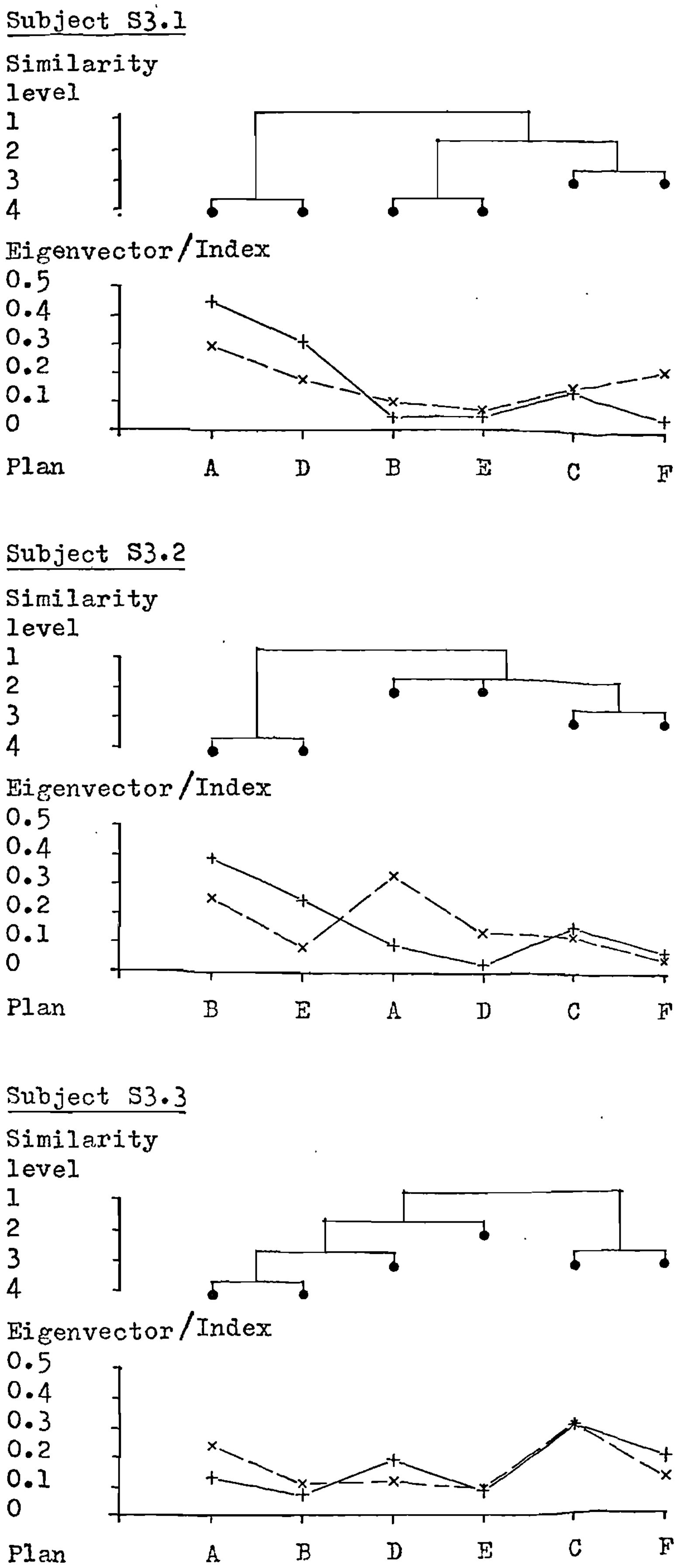
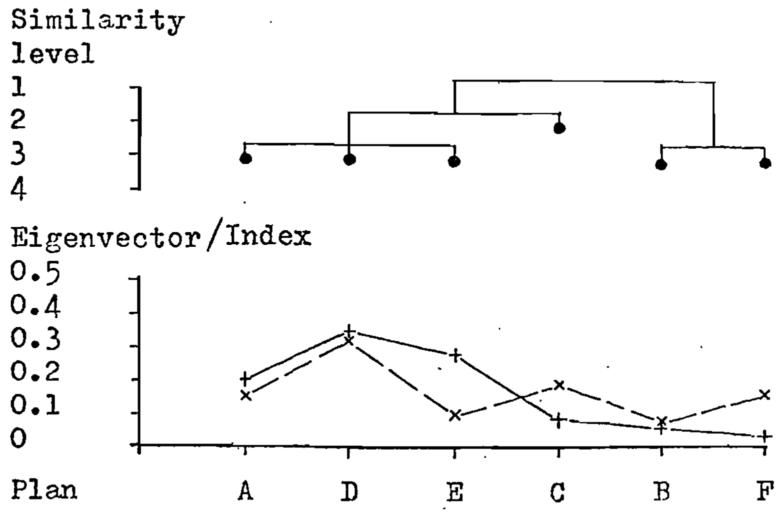
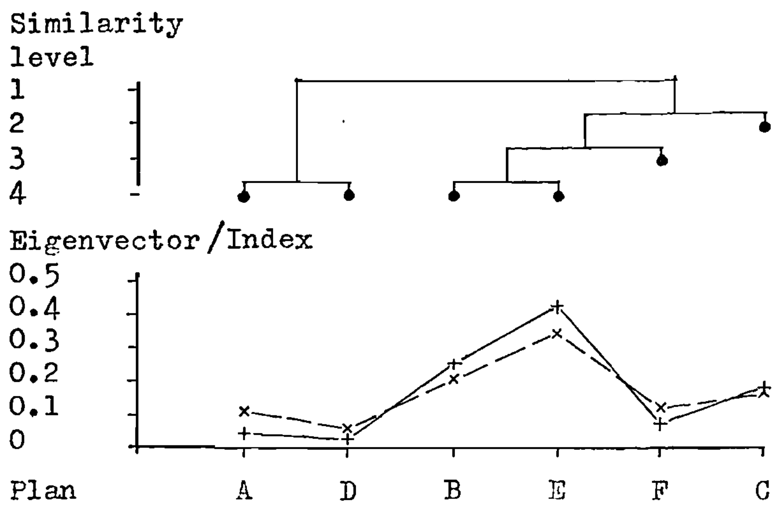


Figure 7.1 continued

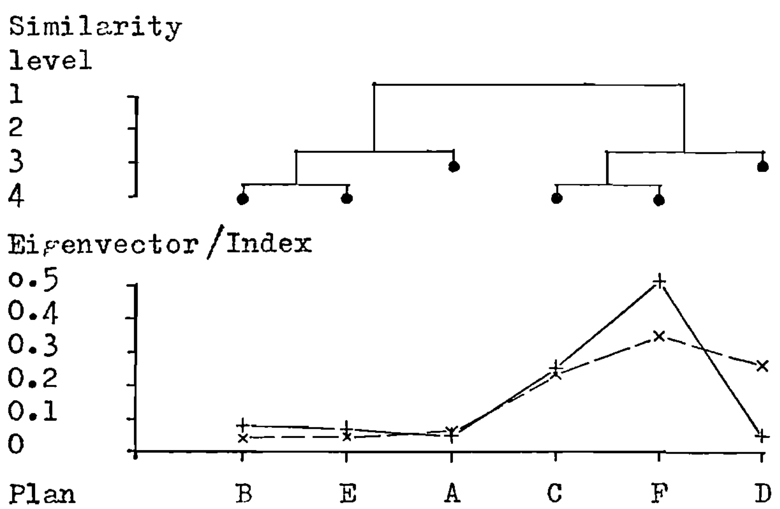
Subject S3.4



Subject S3.5



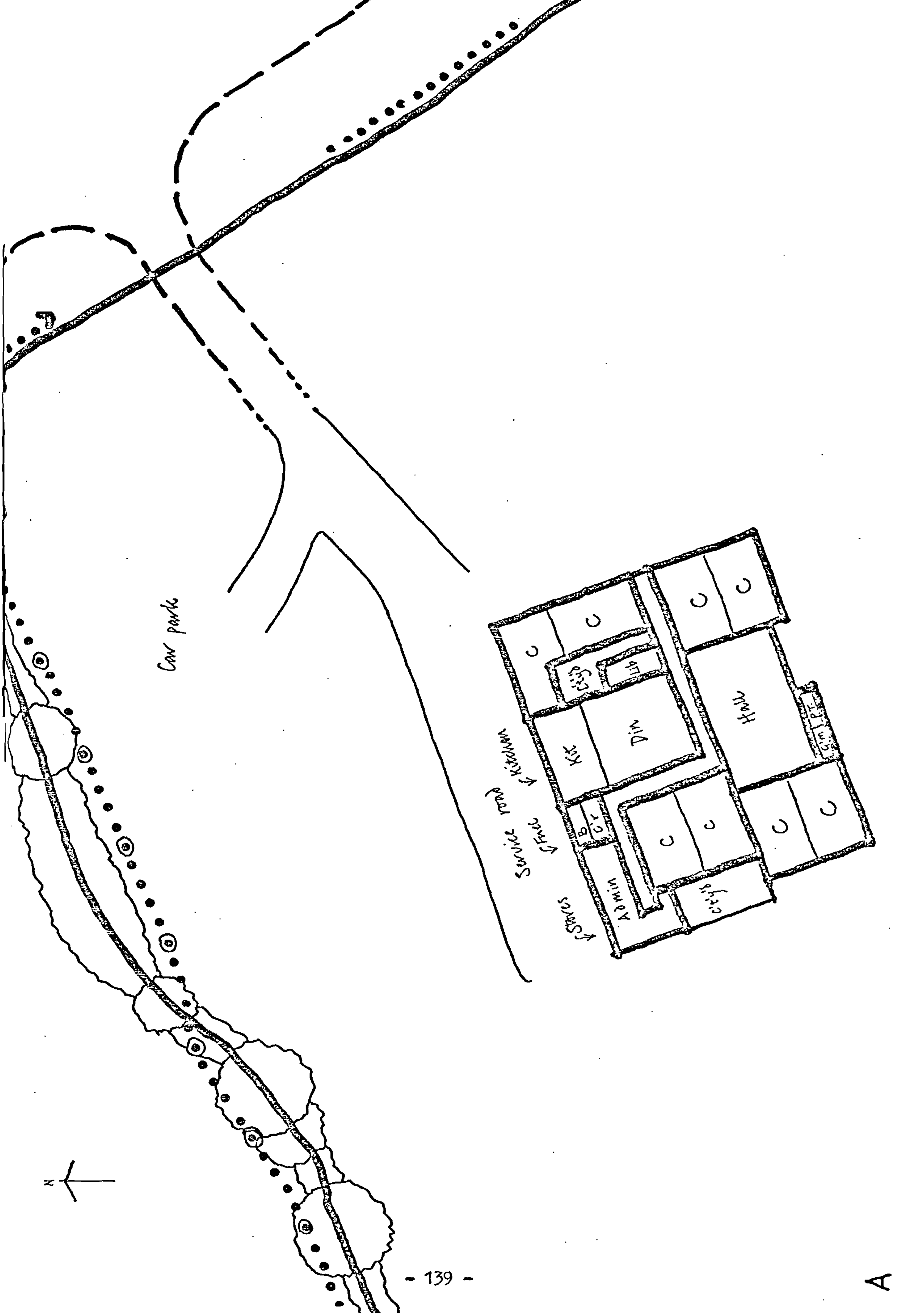
Subject S3.6

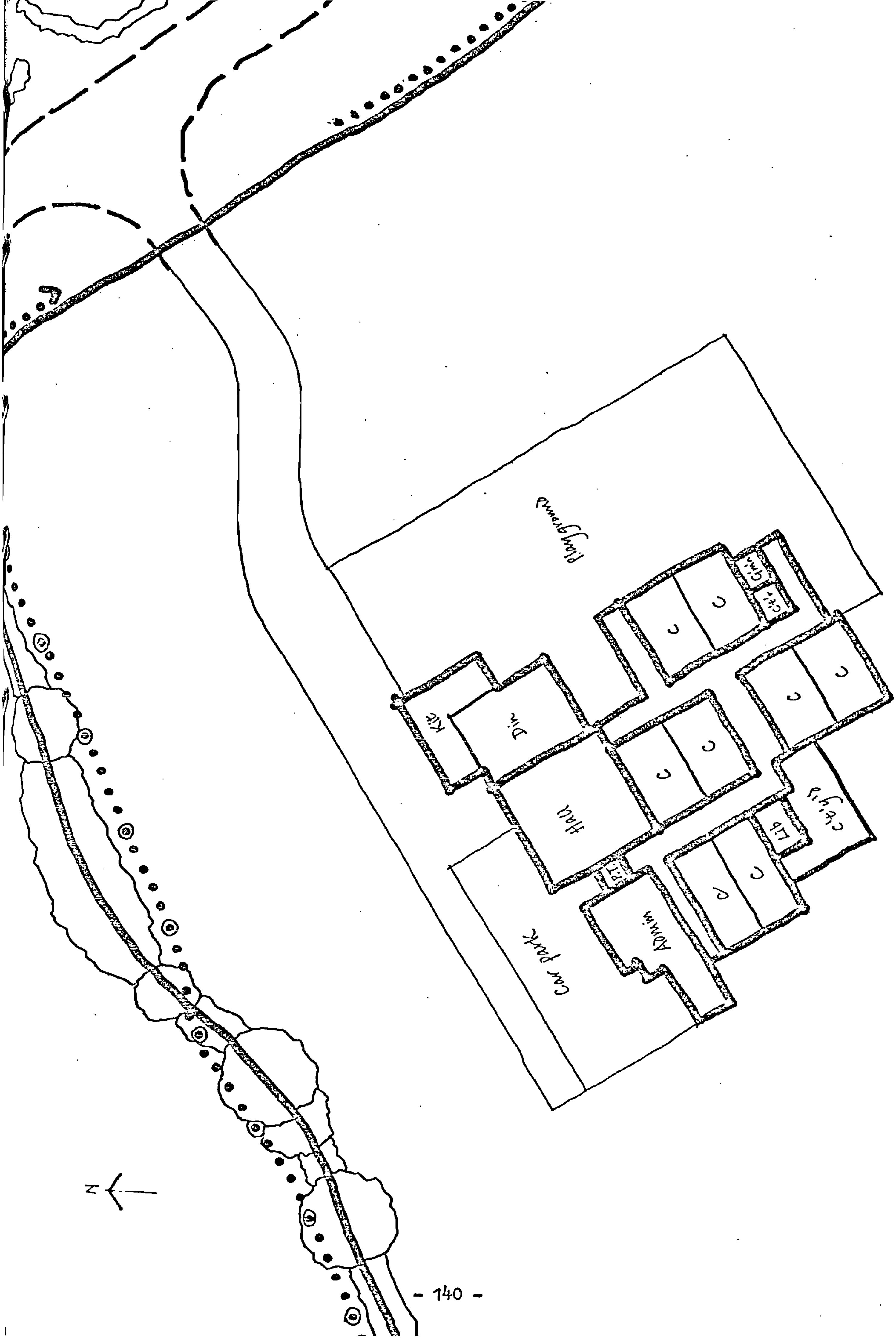


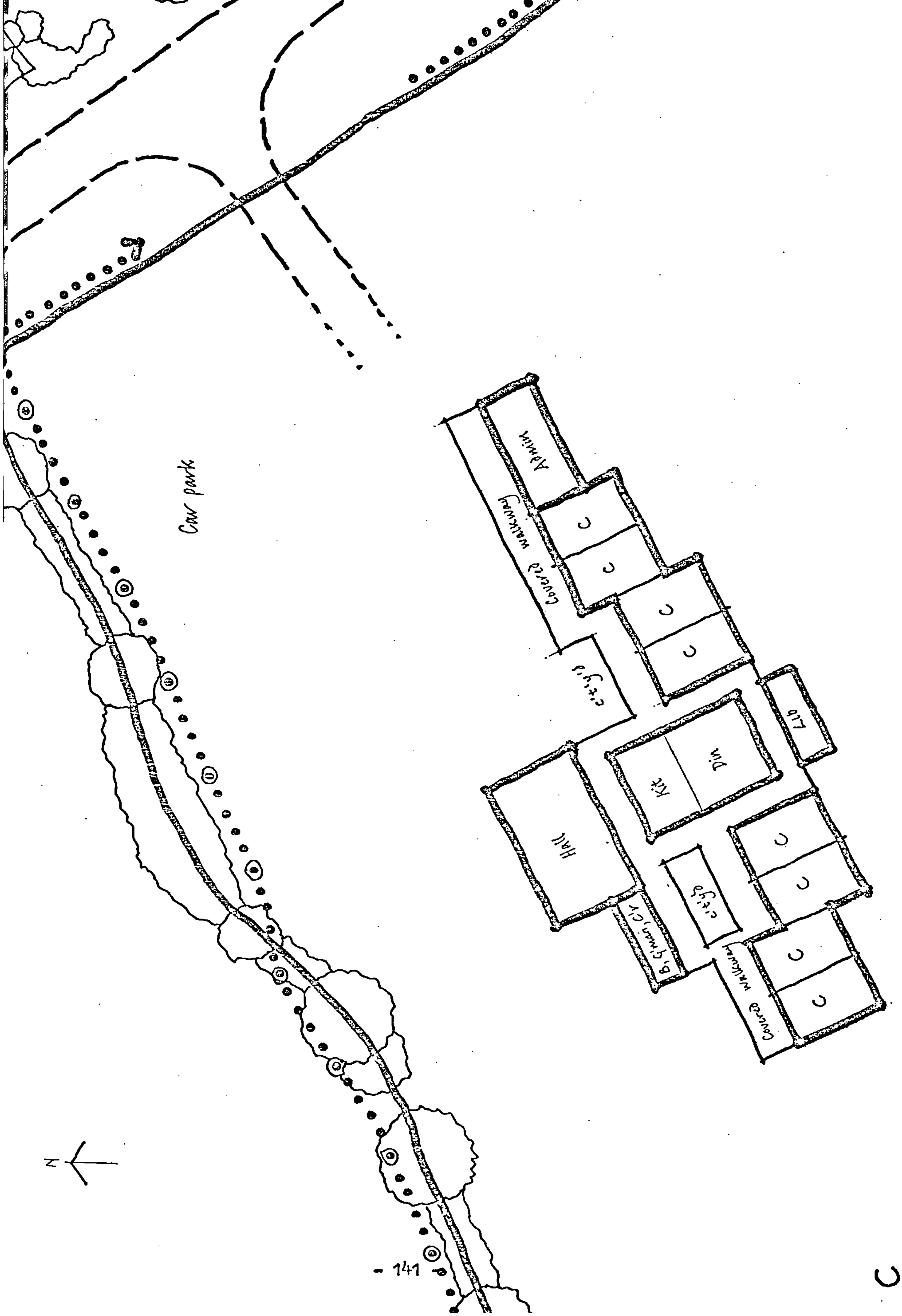
Key +——— Overall evaluation
×– – – Additive utility model indices

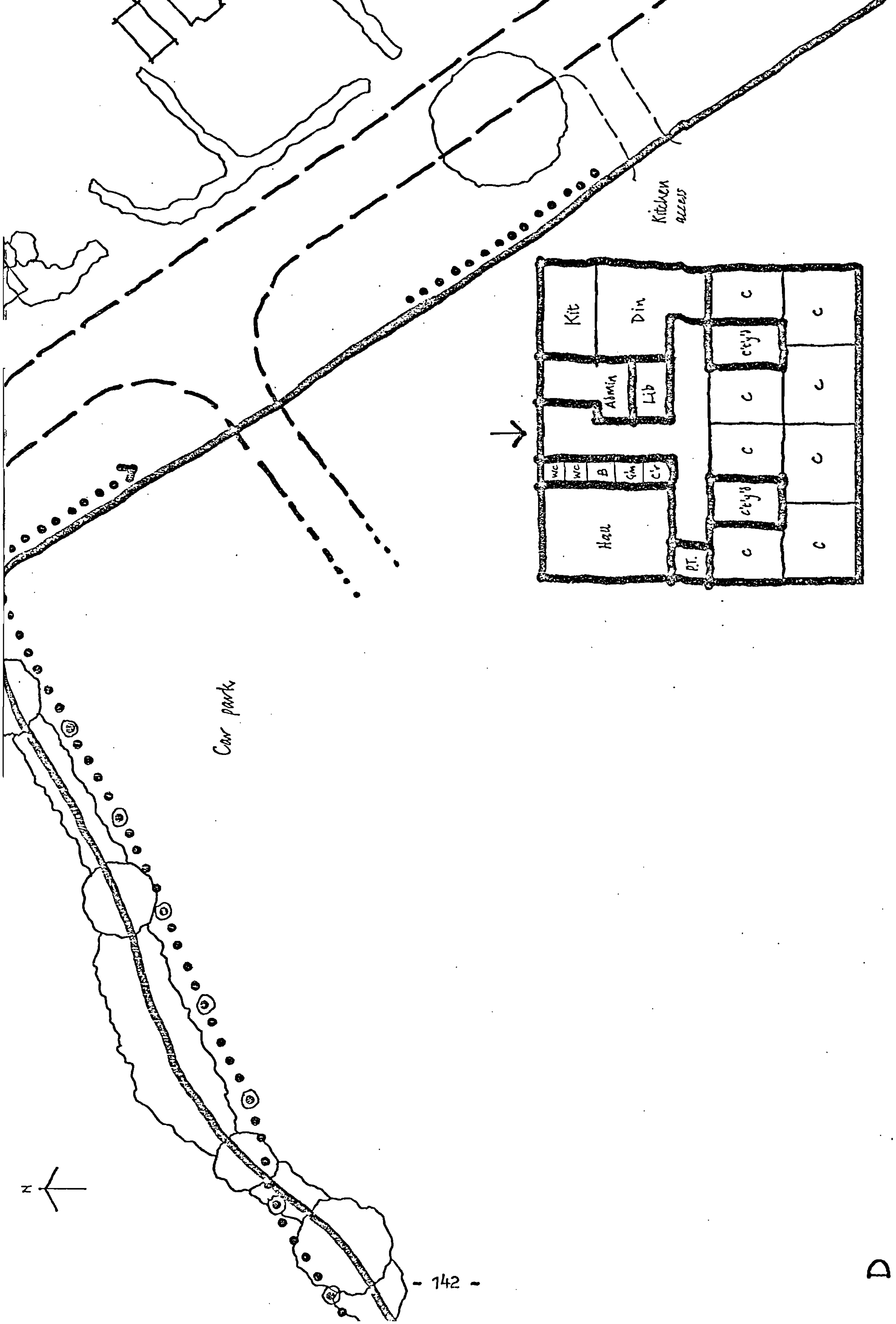
Figure 7.2 The redrawn school plans

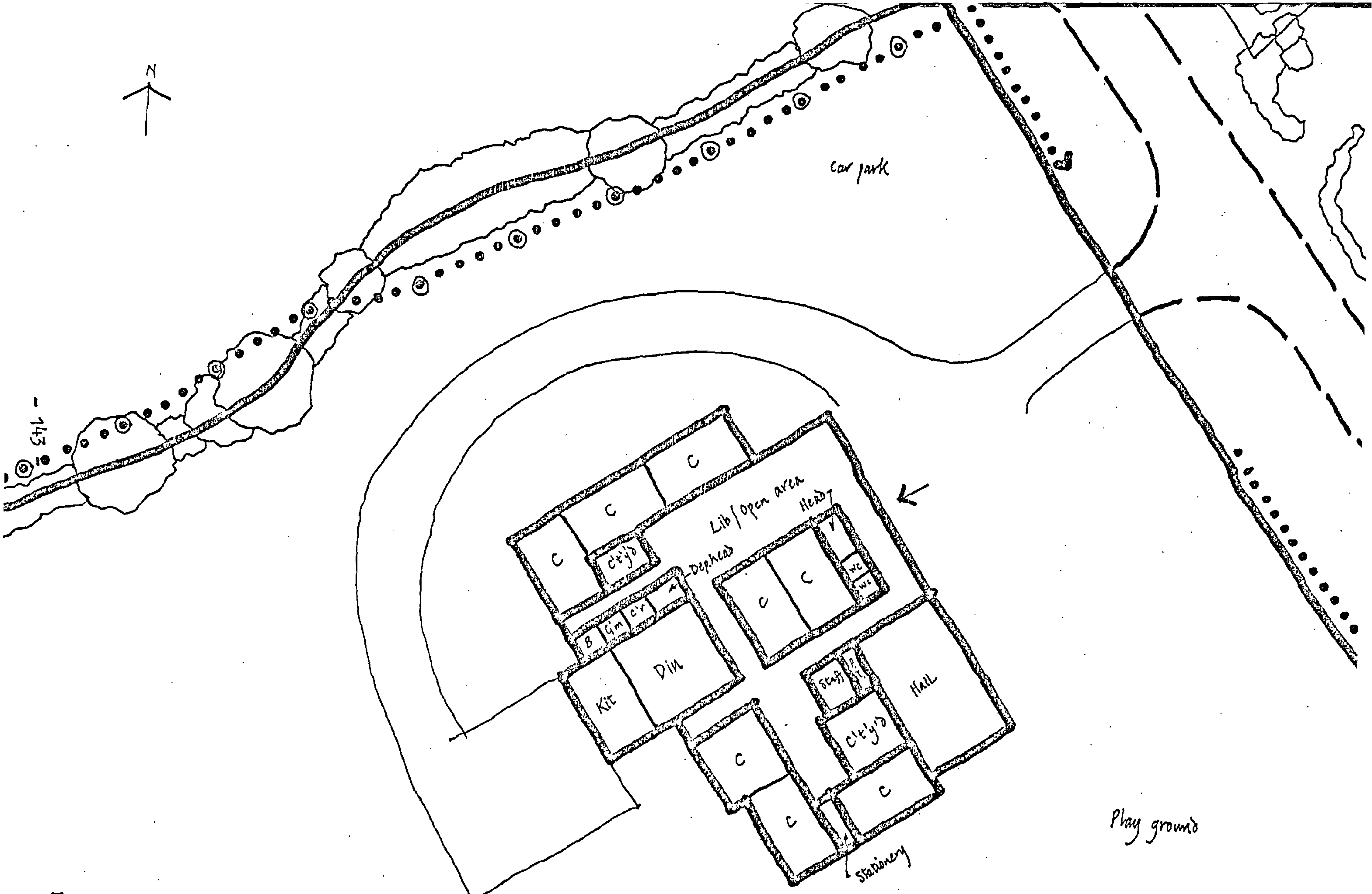
The following six pages show the redrawn school plans prepared in Experiment Three.


Car park
Service road
Fuel
Kitchen
Stores
C
Kit
Din
Hall
Lib
Admin
B
N
A

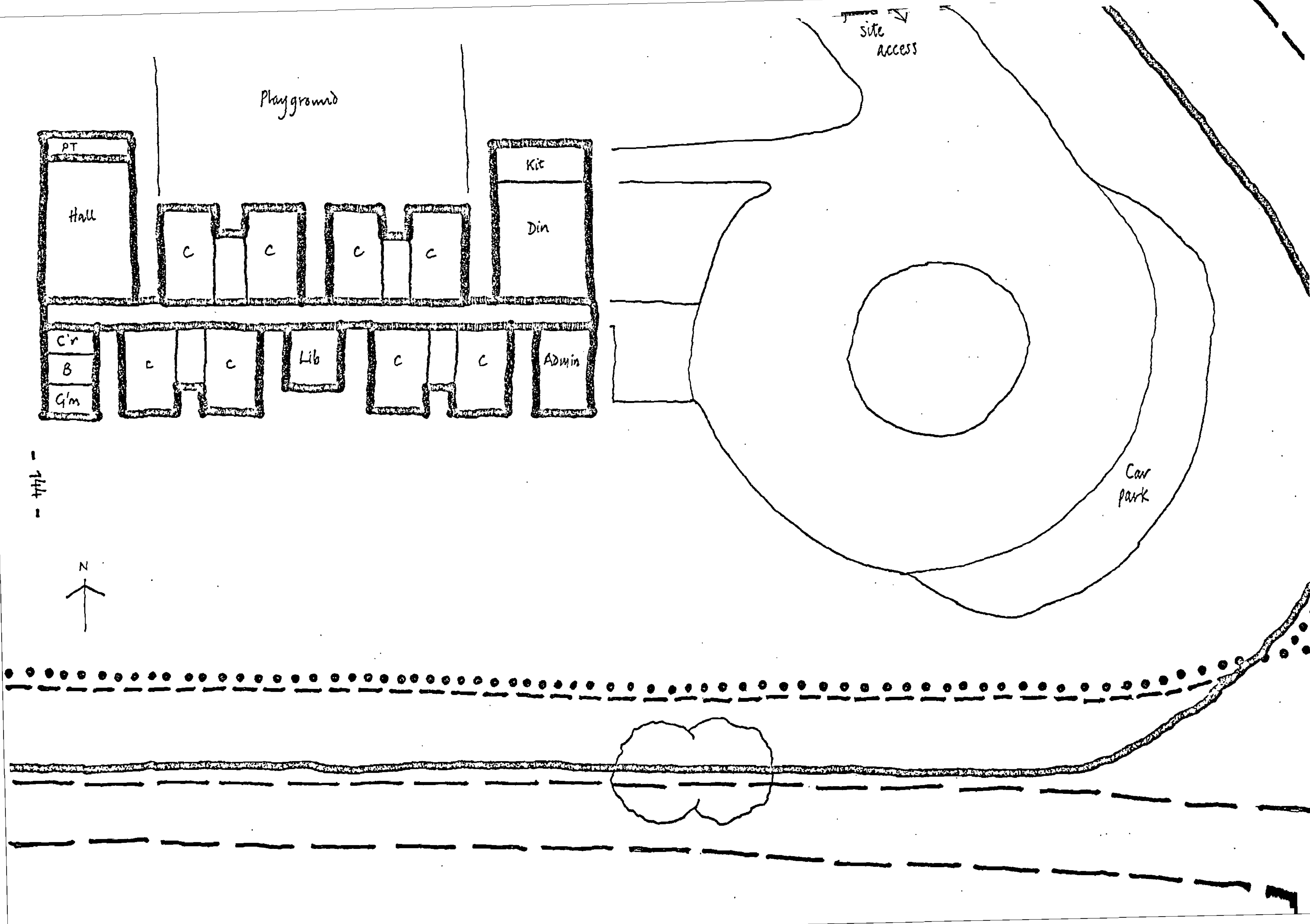

Playground
Kit
Din
Hall
Car park
Admin
C
Lib
N


Car park
Covered walkway
Admin
C
C
C
C
Lib
Kit
Din
Hall
C
C
Covered walkway
C
C
N


Kitchen access
Kit
Din
Admin
Lib
wc
wc
B
Gm
Cr
Hall
P.T.
c'ty'd
c'ty'd
c
Car park
N
D


N
car park
Lib / Open area
Head
Dep.head
C
C
C
C
C
C
C
C
C'tyd
C'tyd
WC
WC
B
Gm
C'r
Din
Kit
Staff
Hall
Stationery
Play ground

E



site access
Playground
PT
Hall
C'r
B
G'm
C
C
C
C
C
C
Lib
C
C
Kit
Din
Admin
Car park
N

CHAPTER 8 EXPERIMENT FOUR

8.1 Introduction

The most important difference between this experiment and Experiment Three is that the subjects taking part were all graduate students of architecture in their first year at the School of Environmental Design, Royal College of Art. This fact apart, the organisation of the experiment virtually repeats that of Experiment Three, with only a few changes in response to the findings of that experiment.

A one-day intensive design exercise was again used as a vehicle for the experiment. The brief and site were both identical to those used in Experiment Three and a similar time-table was followed.

The pattern of the experiment was that attributes for the design of a school plan were elicited from the subjects by brainstorming, and a set of six common attributes agreed through discussion. The attributes were scaled by each subject before and after the design process. Each subject designed a school plan. The plans produced were redrawn by the experimenter. Each subject evaluated all the schemes with respect to each attribute, and for overall merit, and also performed similarity judgements in triadic comparisons of the plans.

The differences made in the organisation of this experiment in response to previous findings may be summarised as follows. First it was felt essential to ensure that the changes being measured in the rating of attributes before and after design were actually caused by the design process rather than occurring arbitrarily. For this reason each subject acted as his own control. In addition to the eliciting of attributes for a school plan, attributes were also elicited for a holiday companion by brainstorming, and a list of six agreed through discussion. These were scaled by each subject before the school plan attributes had been elicited, and then again after design after the school plan attributes had been scaled for the second time. This ensured that the scaling of school plan attributes occurred immediately before and immediately after the design process. The second difference was in direct response to the finding that, while for some subjects the rating of attributes for a school plan changed after the design process, they did not change during evaluation. Therefore the repeat of the scaling of attributes during evaluation was abandoned. A thirddifference was to alter the order in which the subjects performed the evaluations. In this experiment they rated the plans for overall merit after having rated them formally with respect to each of the attributes. It was hoped thatthis might lead to an improvement in the correlation between the additive utility model indices and the overall merit rating.

8.2 Hypotheses

The experiment was designed to test a number of hypotheses. All are expressed as null hypotheses.

8.2.1 Correlations between the ratings of attributes by each subject before and after design

Hypothesis 1 That for each subject there would be no significant correlation between his rating of attributes for a holiday companion before design compared with after design.

Hypothesis 2 That for each subject there would be no significant correlation between his rating of attributes for a school plan before design compared with after design.

8.2.2 Concordance between the subjects' ratings of attributes

Hypothesis 3 That there would be no significant concordance between the subjects' ratings of attributes for a holiday companion before design.

Hypothesis 4 That there would be no significant concordance between the subjects' ratings of attributes for a school plan before design.

Hypothesis 5 That there would be no significant concordance between the subjects' ratings of attributes for a school plan after design.

Hypothesis 6 That there would be no significant concordance between the subjects' ratings of attributes for a holiday companion after design.

8.2.3 The effect of the order in which attributes were scaled

Hypothesis 7 That for each subject there would be no significant correlation between the rating of attributes for a holiday companion before design compared with the rating of attributes for a school plan before design, according to the order in which they were scaled.

Hypothesis 8 That for each subject there would be no significant correlation between the rating of attributes for a school plan after design compared with the rating of attributes for a holiday companion after design, according to the order in which they were scaled.

8.2.4 Differences between the consistency achieved in scaling attributes before and after design

Hypothesis 9 That there would be no significant difference between the set of consistency measures (eigenvalues) achieved in scaling attributes for a holiday companion before design compared with that achieved in scaling attributes for a school plan before design.

Hypothesis 10 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a holiday companion before design compared with that achieved after design.

Hypothesis 11 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a school plan before design compared with that achieved after design.

Hypothesis 12 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a school plan after design compared with that achieved in scaling attributes for a holiday companion after design.

8.2.5 Correlation between the ranks of consistency measures achieved in scaling attributes

Hypothesis 13 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes for a holiday companion before design compared with those achieved in scaling attributes for a school plan before design.

Hypothesis 14 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes for a holiday companion before design compared with after design.

Hypothesis 15 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes for a school plan before design compared with after design.

Hypothesis 16 That there would be no significant correlation between the ranks of consistency measures achieved in scaling attributes for a school plan after design compared with those achieved in scaling attributes for a holiday companion after design.

8.2.6 Concordance between evaluations of school plans

Hypothesis 17 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute a.

Hypothesis 18 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute b.

Hypothesis 19 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute c.

Hypothesis 20 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute d.

Hypothesis 21 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute e.

Hypothesis 22 That there would be no significant concordance between the subjects' ratings of school plans with respect to attribute f.

Hypothesis 23 That there would be no significant concordance between the subjects' ratings of school plans given by the additive utility model indices.

Hypothesis 24 That there would be no significant concordance between the subjects' overall ratings of school plans.

8.2.7 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

Hypothesis 25 That for each subject there would be no significant correlation between the overall rating of school plans and the rating given by additive utility model indices.

Hypothesis 26 That for each subject there would be no correspondence between the overall rating of school plans and the hierarchical cluster analysis of similarity judgements of school plans.

Hypothesis 27 That for each subject there would be no correspondence between the additive utility model indices and the hierarchical cluster analysis of similarity judgements of school plans.

8.3 Experimental Method

In order to test the hypotheses the experiment was organised in a design studio, as follows:

1 The subjects (S4.1, S4.2, S4.3, S4.4, S4.5, S4.6) were given a brief verbal description of how the exercise would be conducted.

2 The subjects, in a group, were read the following statement:

> "Consider the proposition of a fortnight's holiday in Greece. What important attributes or qualities would you look for in the selection of a holiday companion?"

They were asked as a group to 'brainstorm' (Osborn, 1957) to offer spontaneously the attributes they considered important.

3 The subjects were asked to combine or to discard attributes in order to end up with the six which they as a group felt to be the most important in the selection. Through group discussion a list of six common attributes was agreed. The attributes were labelled 'a' to 'f'.

4 Each subject was then asked to work individually in the rating of these six common attributes for a holiday companion. Each was given six blank cards, and was asked to copy the attributes onto the cards, one attribute per card. Each subject was also given the numerical scale to be used (appendix 4.1) and a form with instructions on how to work through the cards and to scale all possible pairs of attributes.
The completed forms were then retained by the experimenter.

5 Then the subjects were given copies of the site plan and brief for a two-form entry primary school. The following statement was read to the group:

> "Consider the implications of planning a primary school, on the given site and to satisfy the given brief. What important attributes or qualities would you take into account in planning the school?"

They were asked as a group to 'brainstorm' to offer spontaneously the attributes they considered important. Two points were emphasised: first that criticism of the attributes was ruled out, second that it was the planning that was important as opposed to, for instance, fittings and finishes. At this session 19 attributes were produced in fifteen minutes. All attributes were recorded on a blackboard, fully visible to the group.

6 The subjects were asked to combine or to discard attributes in order to end up with the six which they as a group felt were the most representative or expressive of the important attributes in planning the school. Through group discussion a list of six common attributes was agreed. The attributes were labelled 'a' to 'f'.

7 Each subject was then asked to work individually in the rating of these six common attributes for a school plan. Each was given six blank cards, and was asked to copy the attributes onto the cards, one attribute per card. Each subject was also given the numerical scale to be used (appendix 4.1) and a form with instructions on how to work through the cards and to scale all possible pairs of attributes (appendix 4.2). The completed forms were then retained by the experimenter.

8 Each subject then proceeded to design a school plan for the given site to satisfy the given brief. Each worked individually.

9 Having designed a school plan, each subject again rated the six attributes for a school plan, using the same technique as he had before design.

10 Each subject rated again the six attributes for a holiday companion, using the same technique as he had before design.

11 The six school plans produced were re-drawn by the experimenter to a standard scale (1:500), orientation and format (figure 8.2).

12 Each subject gave a two-minute presentation of his redrawn school plan to the other five subjects.

13 The redrawn plans were evaluated by the subjects with respect to each of the six attributes. Each subject worked individually. Each was given the six redrawn

plans, the numerical scale to be used and six forms with instructions on how to work through and to scale all possible pairs of plans. Each form stated at the top the name of the attribute with respect to which the plans were to be scaled. The six forms were completed each in turn according to the alphabetical order of the attributes.

14 The redrawn plans were then evaluated by the subjects for their overall merit as schools. Each subject worked individually. Each was given the set of school plans, the numerical scale to be used and a form with instructions on how to work through and to scale all possible pairs of plans.

15 Finally triadic comparisons of the school plans were made. Each subject worked individually, using a form which gave instructions on how to work through all possible triads of plans (appendix 4.5). In differentiating between plans subjects were asked to state the attributes which they had used; they were not constrained that these attributes necessarily be chosen from the six common attributes. Unfortunately S4.4 felt unable to complete this part of the evaluation.

8.4 Results

8.4.1 Correlation between the ratings of attributes by each subject before and after design

Hypotheses 1 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for a holiday companion before and after design. Table 8.2 shows the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S4.1	0.943	.01
S4.2	0.943	.01
S4.3	0.829	.05
S4.4	0.986	.01
S4.5	0.757	NS
S4.6	0.943	.01

In five out of six cases there was significant correlation between the ranks. In the case of S4.5 there was perfect correlation between the before and after design ranks when he had been asked for a simple rank order, but a certain inconsistency in scaling the paired comparisons before design (as shown by the high maximum eigenvalue in table 8.3) resulted in a rank correlation coefficient lower than that which is statistically significant. The null hypothesis was rejected.

Hypothesis 2 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for a school plan before and after design. Table 8.2 shows the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S4.1	0.714	NS
S4.2	0.886	.05
S4.3	0.086	NS
S4.4	0.771	NS
S4.5	0.943	.01
S4.6	0.943	.01

In three out of six cases there was not significant correlation between the ranks; that is, there was a change in the ranking of attributes after design by these subjects. Of these, S4.1's and S4.4's apparent changes could be explained by inconsistency in the paired comparisons; there was almost perfect consistency between the before design and after design ranks when they had been asked for a simple rank order. However for subject S4.3 there was a definite change in the ranking of attributes before and after design. In the other three cases there was significant correlation between the before design and after design ranks of attributes; in these cases there was no change.

8.4.2 Concordance between the subjects' ratings of attributes

Hypotheses 3, 4, 5 and 6 were tested by calculating Kendall's coefficient of concordance (W) between the sets of ranks of attributes under consideration. Table 8.2 shows the data. The following tabulation shows for each hypothesis (H) the set of rankings being tested, the corresponding value of W and the significance level.

H	Ranking of attributes				W	Sig. level
	Before design		After design			
	Holiday companion	School plan	School plan	Holiday companion		
3	x				0.496	.01
4		x			0.508	.01
5			x		0.670	.01
6				x	0.496	.01

Null hypothesis 3 was rejected.
Null hypothesis 4 was rejected.
Null hypothesis 5 was rejected.
Null hypothesis 6 was rejected.

8.4.3 The effect of the order in which attributes were scaled

Hypothesis 7 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for a holiday companion before design and the ranks of attributes for a school plan before design. Table 8.2 shows the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level.

Subject	Correlation coefficient	Sig. level
S4.1	0.314	NS
S4.2	-0.086	NS
S4.3	0.371	NS

Subject	Correlation coefficient	Sig. level
S4.4	0.243	NS
S4.5	0.543	NS
S4.6	0.257	NS

The null hypothesis was not rejected.

Hypothesis 8 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for a school plan after design and the ranks of attributes for a holiday companion after design. Table 8.2 shows the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level.

Subject	Correlation coefficient	Sig. level
S4.1	0.714	NS
S4.2	0.257	NS
S4.3	0.743	NS
S4.4	-0.257	NS
S4.5	0.386	NS
S4.6	0.086	NS

The null hypothesis was not rejected.

8.4.4 Differences between the consistency achieved in scaling attributes before and after design

Hypotheses 9, 10, 11 and 12 were tested by calculating Mann-Whitney's U between the pairs of sets of consistency measures under consideration. Table 8.3 shows the data. The following tabulation shows the pairs of eigenvalues being compared, the corresponding value of U and the probability under the null hypotheses (one-tailed test).

H	Before design		After design		U	Probability under H_o
	Holiday companion	School plan	School plan	Holiday companion		
9	x	x			17	0.469
10	x			x	9	0.090
11		x	x		11	0.155
12			x	x	9	0.090

Null hypothesis 9 was not rejected.
Null hypothesis 10 was not rejected.
Null hypothesis 11 was not rejected.
Null hypothesis 12 was not rejected.

8.4.5 Correlation between the ranks of consistency measures achieved in scaling attributes

Hypotheses 13, 14, 15 and 16 were tested by calculating Spearman's rank correlation coefficient between pairs of sets of ranks of consistency measures. Table 8.3 shows the data. The following tabulation shows for each hypothesis (H) the pairs of sets of ranks being compared, the rank correlation coefficient and the significance level (one-tailed test).

H	Before design		After design		Correlation coefficient	Sig. level
	Holiday companion	School plan	School plan	Holiday companion		
13	x	x			0.029	NS
14	x			x	0.200	NS
15		x	x		0.771	NS
16			x	x	-0.086	NS

Null hypothesis 13 was not rejected.
Null hypothesis 14 was not rejected.
Null hypothesis 15 was not rejected.
Null hypothesis 16 was not rejected.

8.4.6 Concordance between evaluations of school plans

Hypotheses 17 to 22 were tested by calculating Kendall's coefficient of concordance (W) between the sets of ranks under consideration. Table 8.4 shows the data. The following tabulation shows for each hypothesis (H) the sets of ranks being compared, the corresponding value of W and the significance level.

H	Evaluation of plans with respect to attributes						W	Sig. level
	a	b	c	d	e	f		
17	x						0.528	.01
18		x					0.590	.01
19			x				0.489	.01
20				x			0.487	.01
21					x		0.613	.01
22						x	0.702	.01

Null hypothesis 17 was rejected.
Null hypothesis 18 was rejected.
Null hypothesis 19 was rejected.
Null hypothesis 20 was rejected.
Null hypothesis 21 was rejected.
Null hypothesis 22 was rejected.

Hypothesis 23 was tested by calculating Kendall's coefficient of concordance between the sets of ranks of additive utility model indices. Table 8.5 shows the data. The coefficient was found to be 0.356 which was significant ($\alpha = 0.05$). The null hypothesis was rejected.

Hypothesis 24 was tested by calculating Kendall's coefficient of concordance between the sets of ranks of the subjects' overall ratings of school plans. The coefficient was found to be 0.575 which was significant ($\alpha = 0.01$). The null hypothesis was rejected.

8.4.7 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

Hypothesis 25 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of overall ratings of school plans and the ranks of additive utility model indices. Tables 8.5 and 8.6 show the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S4.1	0.829	.05
S4.2	0.943	.01
S4.3	0.029	NS
S4.4	-0.029	NS
S4.5	1.000	.01
S4.6	-0.771	NS

The null hypothesis was not entirely rejected.

Hypotheses 26 and 27 were tested by comparing for each subject the hierarchical cluster analysis of similarity judgements with both the overall ratings of school plans and the additive utility model indices. Figure 8.1 shows the data. The following tabulation gives a verbal estimate on a four point scale, 'high', 'medium', 'low', 'no', of the correspondence, for each subject, between the hierarchical cluster analysis, the overall rating and the indices.

Subject	Correspondence with hierarchical cluster analysis	
	Overall rating	Index
S4.1	High correspondence. Cluster E-F: both rated close. Cluster A-B: both rated high. D-C: rated low.	High correspondence. Cluster E-F: both rated close. Cluster A-B: both rated high. D-C: rated low.
S4.2	High correspondence. Cluster E-F: both rated close. Cluster A-C-D: all rated low. Plan B: rated highly.	High correspondence. Cluster E-F: both rated close. Cluster A-C-D: all rated low. Plan B: rated highly.
S4.3	Medium correspondence. Cluster E-F and A: all rated close. Plan C: rated low.	Low correspondence. Cluster E-F and A: all rated close.
S4.4	Subject did not perform similarity judgements.	
S4.5	High correspondence. Cluster E-F: both rated highly. Plan C: rated low.	High correspondence. Cluster E-F: both rated highly. Plan C: rated low.
S4.6	Medium correspondence. Cluster E-F: both rated close. Cluster B-D: both rated close.	Medium correspondence. Cluster E-F: both rated close. Cluster B-D: both rated close.

Null hypothesis 26 was rejected.
Null hypothesis 27 was rejected.

8.5 Conclusions

8.5.1 The effect of the design process on rating attributes

The scaling of attributes for a holiday companion meant that each subject acted as his own control. Thus changes in the rating of attributes for a school plan by each subject may be checked to see whether they are an arbitrary change for no apparent reason or whether they may be ascribed definitely to the design process. The results of testing hypothesis 1 are therefore most encouraging in terms of the experimental method. In all six subjects' cases (but see 8.4.1) there was significant correlation between the ranks of attributes for a holiday companion before and after design. That is, there were no (significant) arbitrary changes in the ranks of attributes by any subject. In turn this means that if significant changes are found in the attributes for school plan before and after design these changes may be ascribed to the design process.

The result of testing hypothesis 2 takes on considerable importance. In three out of six cases, S4.1, S4.3 and S4.4, there were significant changes in the ratings of attributes before and after design, as recorded by the paired comparisons. For two subjects these recorded changes appear to be due to inconsistency in the scaling of attributes. However for subject S4.3 the design process undoubtedly resulted in a measurable change in his subjective priorities. The other three designers, S4.2, S4.5 and S4.6, conversely, had a relatively fixed sense of priorities which did not change significantly as a result of the design process.

8.5.2 Concordance between the subjects' ratings of attributes

The results of testing hypothesis 3 showed that there was significant concordance between the subjects' ranks of attributes for a holiday companion before design. Not surprisingly, as there were no significant changes after the design process there was still significant agreement between the ranks after design. This is an unexpected result.

The results of testing hypothesis 4 showed that there was significant concordance between the subjects' ranks of attributes for a school plan before design. And although three subjects changed their rating of attributes as a result of the design process (but see 8.4.1) there was still highly significant concordance between the ranks after design. The inference is that this group of designers shares a similar set of priorities.

8.5.3 The effect of the order in which attributes were scaled

The results of testing hypotheses 7 and 8 were as expected. They show that the order in which the pairs of attributes were scaled by the subjects had no effect on the scaled judgements.

8.5.4 The effect of the design process on consistency

The results of testing hypotheses 9, 10, 11 and 12 give rise to two conclusions. First it seems clear that successive attempts at performing prioritization do not, of themselves, result in improvements in internal consistency. Second they show that there was no improvement in internal consistency in scaling attributes for a school plan caused by the design process. Only if there had been, would it have been necessary to check that there was no improvement in the control part of the experiment, where holiday companion attributes were scaled.

8.5.5 The ranks of consistency measures in scaling attributes

The results of testing hypotheses 13, 14, 15 and 16 show that personality factors did not influence the consistency achieved in scaling attributes. Those who were most consistent in scaling attributes for a school plan were not necessarily more consistent in scaling attributes for a holiday companion. Neither were those more consistent in scaling attributes for a school plan before design necessarily more consistent after design. There is no reason to believe, therefore, that some subjects were inherently more consistent than others in scaling paired comparisons of attributes. This inference is most encouraging in these experiments because, had the converse been found, comparisons of subjects' consistency measures would have revealed only their inherent consistency as judges rather than the differing degrees of consistency they achieve in scaling paired comparisons of attributes for various purposes. Extending this argument it may be inferred that the consistency measure reveals more than merely a trait of personality; it reveals the judge's ability to achieve consistency for the purpose, and on the occasion, in question.

8.5.6 Concordance between evaluations of school plans

The results of testing hypotheses 17 to 22 were as expected. There was a high level of agreement between the evaluations of alternative designs with respect to individual attributes. This implies that, in contrast with the results in Experiment Three, the evaluators were in agreement as to the planning implications of each verbally stated attribute.

Part of the tentative theory proposed in chapter 2 expressed the expectation that there would not be concordance among evaluations of alternatives for overall merit because the evaluators would not be in agreement about the relative importance of attributes. In the present experiment there is concordance between the subjects' ratings of priorities, and therefore it would be expected that, when evaluating alternative designs, there would be concordance between the evaluations. This proved to be the case; there was significant concordance between the subjects' ratings of school plans for overall merit. Interestingly too, there was concordance between the evaluations given by the additive utility model indices, even though not all the subjects achieved significant correlation between their overall merit ratings of school plans, and the indices.

8.5.7 Overall rating of school plans, additive utility model indices, and hierarchical cluster analysis

The conclusions which may be drawn from the results of testing hypotheses 25, 26 and 27 concern the comparison between three methods of evaluation of the same stimulus items, the school plans. The following tabulation is a combination of the results of testing hypotheses 25, 26 and 27. For each of the three comparisons between the overall merit ratings, the additive utility model indices and the hierarchical cluster analysis, each subject is given a verbal estimate on a four point scale, 'high', 'medium', 'low', 'no', of the degree of correspondence. In this tabulation the rank correlation coefficients found in testing hypothesis 25 are converted to the verbal description thus: correlation significant at 0.01 level, 'high'; correlation significant at 0.05 level, 'medium'; positive correlation, 'low'; negative correlation, 'no'.

	Degree of correspondence between pairs of results		
Overall rating	x	x	
Indices	x		x
Cluster analysis		x	x
S4.1	medium	high	high
S4.2	high	high	high
S4.3	low	medium	low
S4.4	no	-	-
S4.5	high	high	high
S4.6	no	medium	medium

In this tabulation it takes only one set of results to fail to correspond with the others for two of the three pairs to be affected. The results are therefore encouraging in that for all subjects (except S4.4) there is at least one pair of results that shows a high or medium degree of correspondence.

Discrepancies in the five sets of results available occur for subjects S4.3 and S4.6. For subject S4.3 several reasons may be identified. In the triadic comparisons he used only attributes 'a', 'b', 'c', 'f' to differentiate between plans, but did not include attribute 'd' which he had in fact rated the most important. Consistency in scaling attributes even after design was relatively poor, with a maximum eigenvalue of 6.944. Subject S4.6 on the other hand did use all six attributes to differentiate between plans in the triadic comparisons, and achieved good consistency in scaling these attributes after design. What appears to have happened in the computation of the indices is for differences between the ratings of plans to have been almost cancelled; the ratings are approximately equal. The negative correlation coefficient found in testing hypothesis 25 is not in this case a particularly relevant result; figure 8.1 shows much more clearly the pattern of the results.

Much more positive are the results for subjects S4.1, S4.2 and S4.5. For all three subjects, again as figure 8.1 shows clearly, there was definite correspondence between all three sets of results. In particular it may be claimed that for these three subjects the additive utility model provides a precise way of explaining their preferences in terms of the weighting of attributes and the weighting of plans with respect to each attribute.

Table 8.1 Description of attributes

Purpose	Attribute	Description
Holiday companion	a	Compatible sense of humour
	b	Similar attitude towards the sun
	c	Similar financial resources
	d	Punctuality
	e	Willing to be led
	f	Similar interests
School planning	a	Small scale, friendly, non-institutional, easily understood
	b	Efficient, circulation
	c	Orientation
	d	Architectural form
	e	Compactness, capital and running costs
	f	Classroom arrangement; flexibility and outdoor area for each class

Table 8.2 Rating of attributes for a school plan and for a holiday companion before and after design, and coefficient of concordance between judges

	Attributes	Eigenvectors						Ranks						Coefficient of concordance	
		S4.1	S4.2	S4.3	S4.4	S4.5	S4.6	S4.1	S4.2	S4.3	S4.4	S4.5	S4.6	W	Significance level
Holiday companion 1	a	.330	.399	.313	.125	.342	.228	2	1	1	3	1	2	.496	.01
	b	.104	.283	.092	.116	.026	.070	3	2	5	5	6	4		
	c	.097	.052	.077	.119	.179	.161	4	5	6	4	3	3		
	d	.055	.030	.231	.036	.032	.042	5	6	2	6	5	6		
	e	.054	.070	.159	.302	.176	.069	6	4	3	1.5	4	5		
	f	.361	.166	.128	.302	.245	.430	1	3	4	1.5	2	1		
School planning 1	a	.231	.271	.266	.121	.333	.390	3	2	1	3	1	1	.508	.01
	b	.094	.045	.119	.052	.046	.042	4	6	4	5	5	6		
	c	.031	.300	.114	.032	.117	.109	6	1	5	6	4	4		
	d	.308	.147	.105	.318	.295	.230	1	4	6	2	2	2		
	e	.063	.060	.133	.060	.040	.069	5	5	3	4	6	5		
	f	.274	.176	.262	.418	.168	.160	2	3	2	1	3	3		
School planning 2	a	.270	.225	.212	.195	.375	.311	1	3	2	2	1	2	.670	.01
	b	.160	.079	.066	.068	.041	.036	4	5	6	4	6	6		
	c	.059	.263	.085	.028	.110	.090	5	1	5	6	4	4		
	d	.224	.159	.231	.527	.273	.392	3	4	1	1	2	1		
	e	.044	.046	.193	.042	.047	.065	6	6	4	5	5	5		
	f	.243	.228	.212	.141	.154	.106	2	2	3	3	3	3		
Holiday companion 2	a	.334	.362	.313	.155	.294	.255	2	1	1	3	1.5	2	.496	.01
	b	.079	.283	.089	.067	.024	.074	4	2	4	5	6	5		
	c	.090	.083	.073	.078	.087	.188	3	4	6	4	4	3		
	d	.033	.038	.289	.035	.059	.042	5	6	2	6	5	6		
	e	.025	.077	.088	.265	.294	.075	6	5	5	2	1.5	4		
	f	.439	.157	.149	.400	.243	.367	1	3	3	1	3	1		

Table 8.3 Consistency achieved in scaling attributes

Subject	Before design				After design			
	Holiday companion		School planning		School planning		Holiday companion	
	λmax	rank	λmax	rank	λmax	rank	λmax	rank
S4.1	6.242	1	7.306	5	6.434	3	6.425	5
S4.2	6.457	2	6.449	2	6.378	1	6.355	3
S4.3	6.591	4	7.745	6	6.944	6	6.315	1
S4.4	6.941	5	6.901	4	6.545	5	6.404	4
S4.5	7.333	6	6.458	3	6.467	4	6.716	6
S4.6	6.590	3	6.427	1	6.396	2	6.339	2

Table 8.4 Rating of school plans with respect to attributes

Attribute	Plan	Subject											
		Eigenvectors						Ranks					
		S4.1	S4.2	S4.3	S4.4	S4.5	S4.6	S4.1	S4.2	S4.3	S4.4	S4.5	S4.6
a	A	.222	.096	.277	.349	.149	.067	2	4	2	1	3	6
	B	.191	.418	.091	.168	.098	.159	3	1	4	3	4	2.5
	C	.058	.052	.082	.107	.058	.380	5	5	5	5	6	1
	D	.026	.051	.064	.028	.077	.091	6	6	6	6	5	5
	E	.410	.253	.296	.206	.369	.144	1	2	1	2	1	4
	F	.093	.131	.190	.143	.249	.159	4	3	3	4	2	2.5
b	A	.330	.108	.151	.224	.136	.067	2	4	4	1	4	5
	B	.348	.190	.076	.137	.224	.395	1	3	5	4	3	1
	C	.020	.045	.218	.113	.042	.108	6	6	3	5	6	4
	D	.047	.052	.034	.106	.079	.048	5	5	6	6	5	6
	E	.150	.276	.266	.217	.284	.186	3	2	1	2	1	3
	F	.105	.329	.255	.203	.235	.195	4	1	2	3	2	2
c	A	.276	.040	.181	.082	.059	.135	2	6	3	4	5	4
	B	.061	.362	.105	.408	.107	.051	5	1	6	1	4	5
	C	.023	.076	.111	.024	.043	.049	6	5	5	6	6	6
	D	.437	.232	.193	.314	.446	.249	1	2	2	2	1	3
	E	.094	.120	.179	.105	.161	.263	4	4	4	3	3	1
	F	.108	.171	.231	.067	.184	.255	3	3	1	5	2	2
d	A	.529	.190	.114	.274	.270	.199	1	3	4	2	2	2.5
	B	.207	.337	.373	.354	.363	.199	2	1	1	1	1	2.5
	C	.024	.068	.156	.019	.041	.199	6	5	3	6	6	2.5
	D	.143	.067	.060	.131	.146	.199	3	6	6	4	3	2.5
	E	.034	.127	.066	.170	.105	.090	5	4	5	3	4	6
	F	.062	.212	.232	.053	.074	.114	4	2	2	5	5	5
e	A	.183	.053	.035	.030	.071	.096	3	5	5	5	5	5
	B	.446	.133	.218	.071	.112	.193	1	4	2	4	4	2
	C	.032	.361	.535	.538	.336	.288	6	1	1	1	1	1
	D	.034	.039	.029	.027	.035	.073	5	6	6	6	6	6
	E	.198	.166	.085	.137	.255	.178	2	3	4	3	2	3
	F	.108	.249	.098	.197	.192	.173	4	2	3	2	3	4
f	A	.220	.157	.270	.049	.101	.224	3	4	2	5	4	3
	B	.121	.225	.041	.050	.137	.099	4	3	5	4	3	5
	C	.043	.031	.037	.022	.037	.046	5	6	6	6	6	6
	D	.033	.067	.089	.488	.068	.109	6	5	4	1	5	4
	E	.233	.237	.340	.208	.419	.249	2	2	1	2	1	2
	F	.349	.283	.224	.182	.238	.272	1	1	3	3	2	1

Table 8.5 Rating of school plans given by additive utility model indices, corresponding ranks and coefficient of concordance (W)

Plan	Indices						Ranks						W	Significance level
	S4.1	S4.2	S4.3	S4.4	S4.5	S4.6	S4.1	S4.2	S4.3	S4.4	S4.5	S4.6		
A	.309	.109	.175	.238	.161	.143	1	4	4	2	4	6		
B	.206	.315	.170	.250	.183	.169	3	1	5	1	3	2		
C	.038	.070	.188	.064	.061	.228	6	6	3	6	6	1	0.356	.05
D	.082	.104	.070	.160	.133	.147	5	5	6	4	5	5		
E	.213	.192	.199	.183	.273	.148	2	3	1	3	1	4		
F	.152	.210	.197	.105	.189	.164	4	2	2	5	2	3		

Table 8.6 Overall rating of school plans: eigenvectors, corresponding ranks and coefficient of concordance (W)

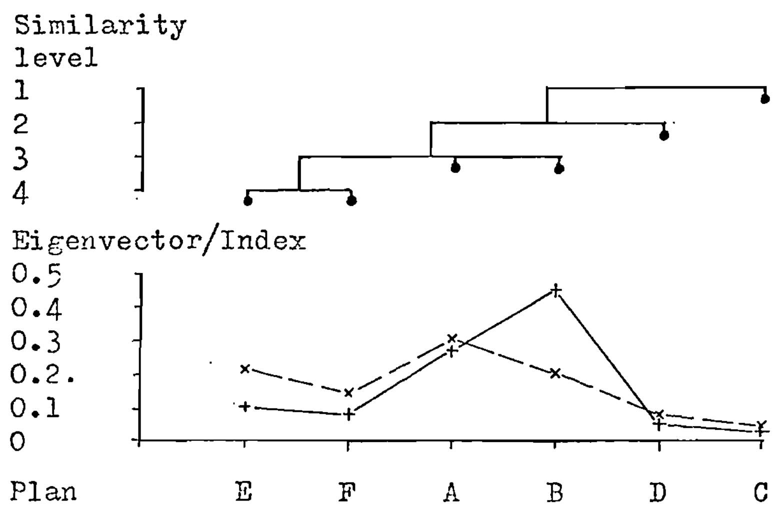
Plan	Eigenvectors						Ranks						W	Significance level
	S4.1	S4.2	S4.3	S4.4	S4.5	S4.6	S4.1	S4.2	S4.3	S4.4	S4.5	S4.6		
A	.270	.153	.219	.045	.101	.308	2	4	2	5	4	1	0.575	.01
B	.459	.288	.504	.084	.164	.125	1	1	1	4	3	5		
C	.024	.032	.043	.028	.029	.060	6	6	5	6	6	6		
D	.055	.048	.041	.125	.051	.130	5	5	6	3	5	4		
E	.104	.251	.083	.236	.412	.156	3	2	4	2	1	3		
F	.088	.228	.110	.483	.243	.221	4	3	3	1	2	2		

Table 8.7 Consistency achieved in scaling plans with respect to attributes

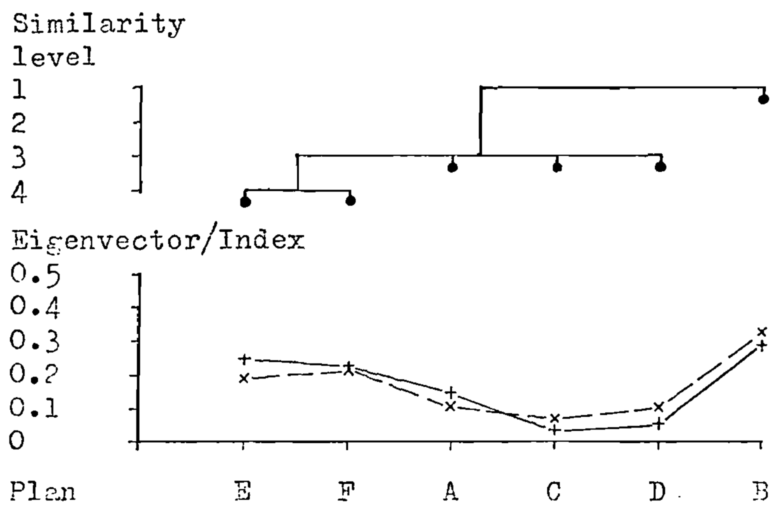
Subject	Attribute					
	λmax					
	a	b	c	d	e	f
S4.1	7.588	7.078	6.874	6.900	7.332	7.164
S4.2	6.141	6.297	6.307	6.119	6.276	6.245
S4.3	7.546	6.222	6.806	6.887	6.496	6.439
S4.4	7.576	7.080	6.376	7.280	6.765	7.002
S4.5	6.288	6.385	6.693	6.560	6.249	6.634
S4.6	6.113	6.602	6.154	6.055	6.184	6.245

Figure 8.1 Comparison between each subject's similarity judgements of school plans, the weighting given in overall evaluation and the additive utility model indices.

Subject S4.1



Subject S4.2



Subject S4.3

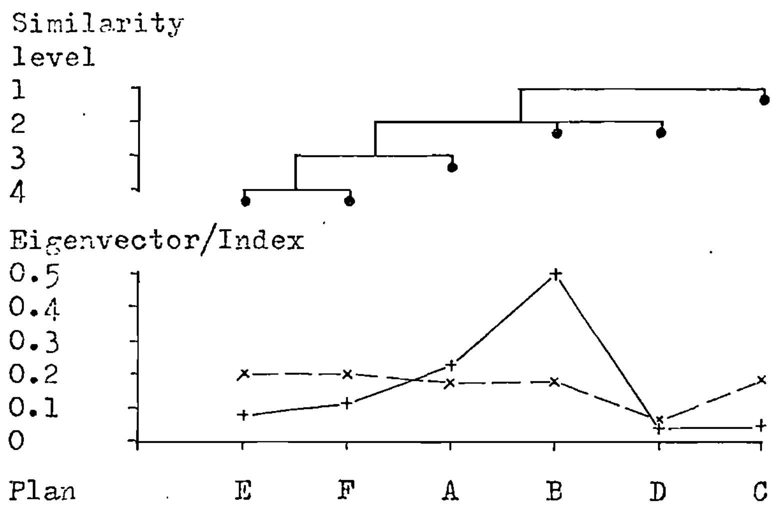
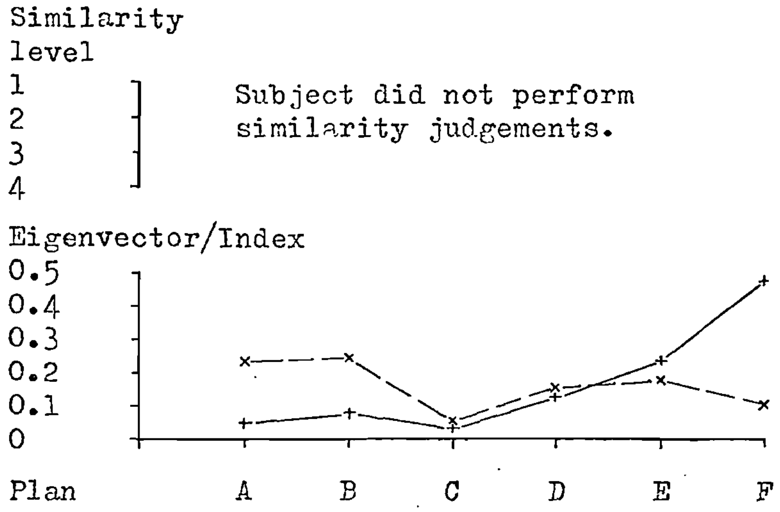
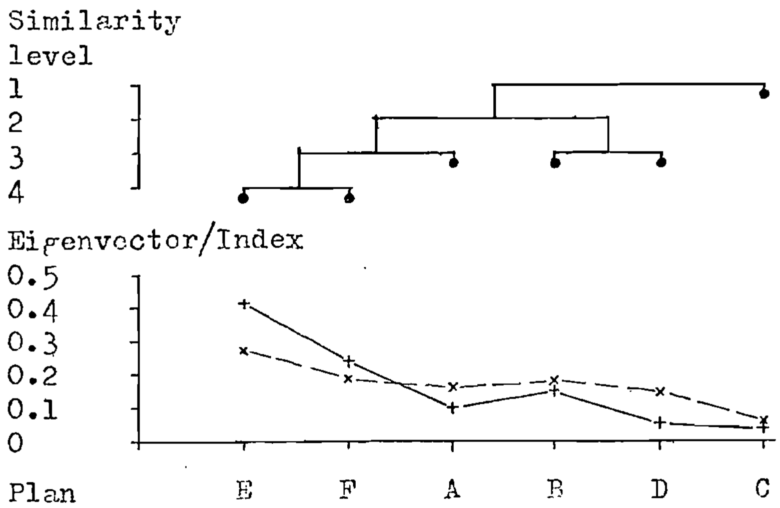


Figure 8.1 continued

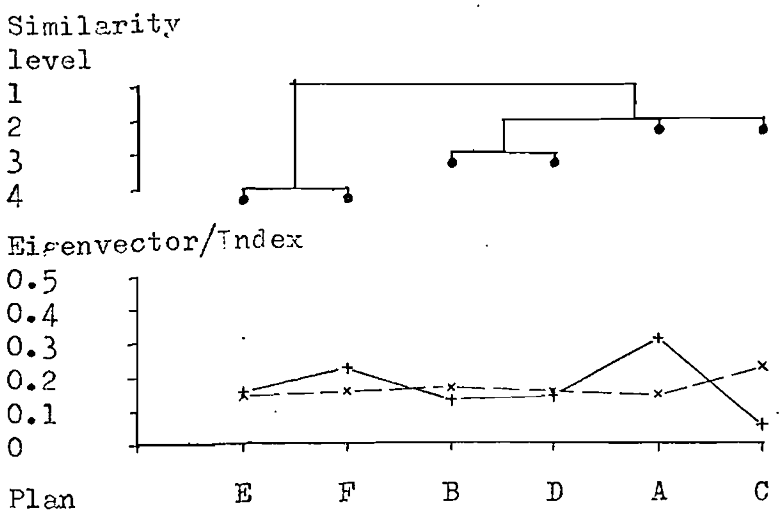
Subject S4.4



Subject S4.5



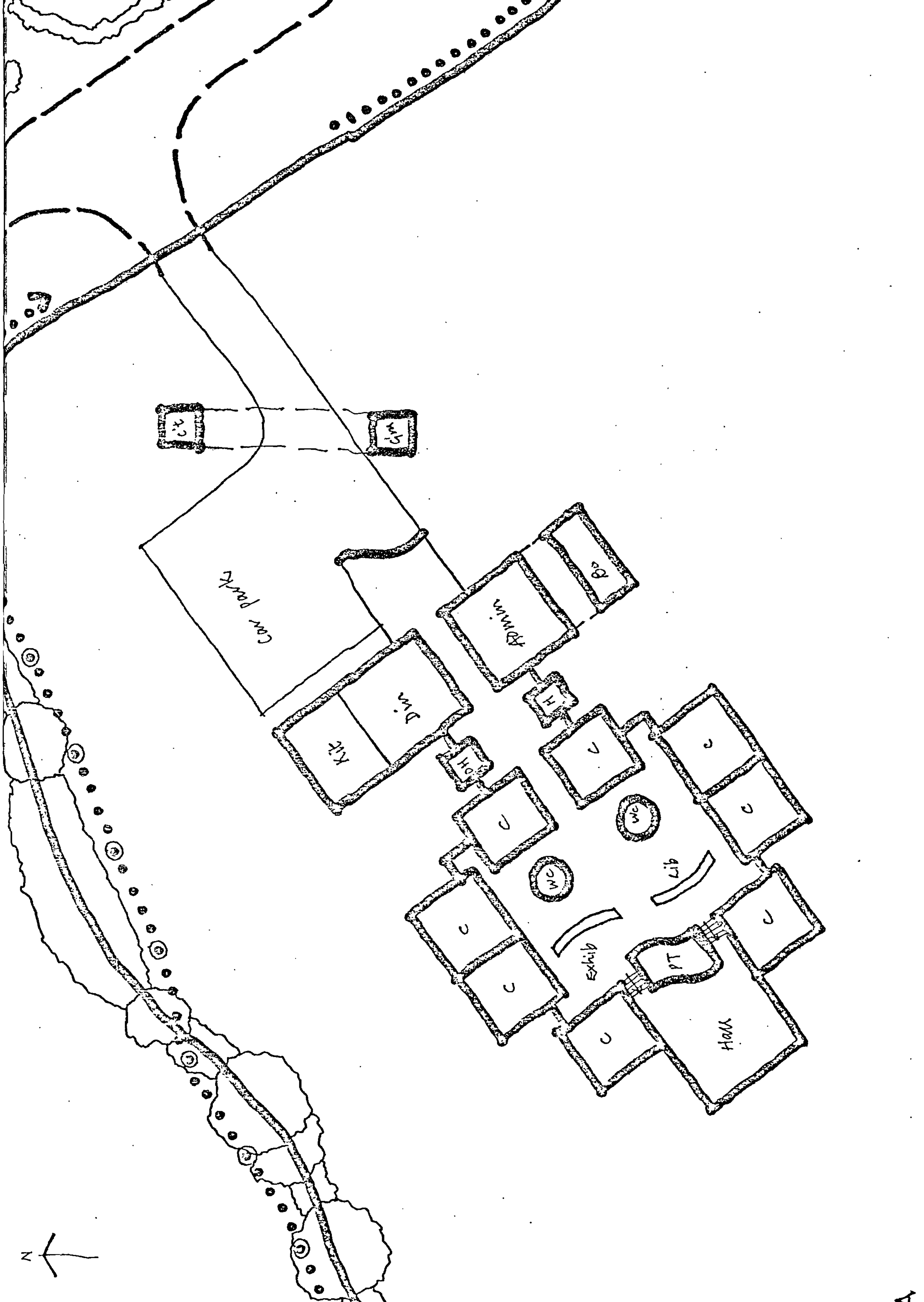
Subject S4.6



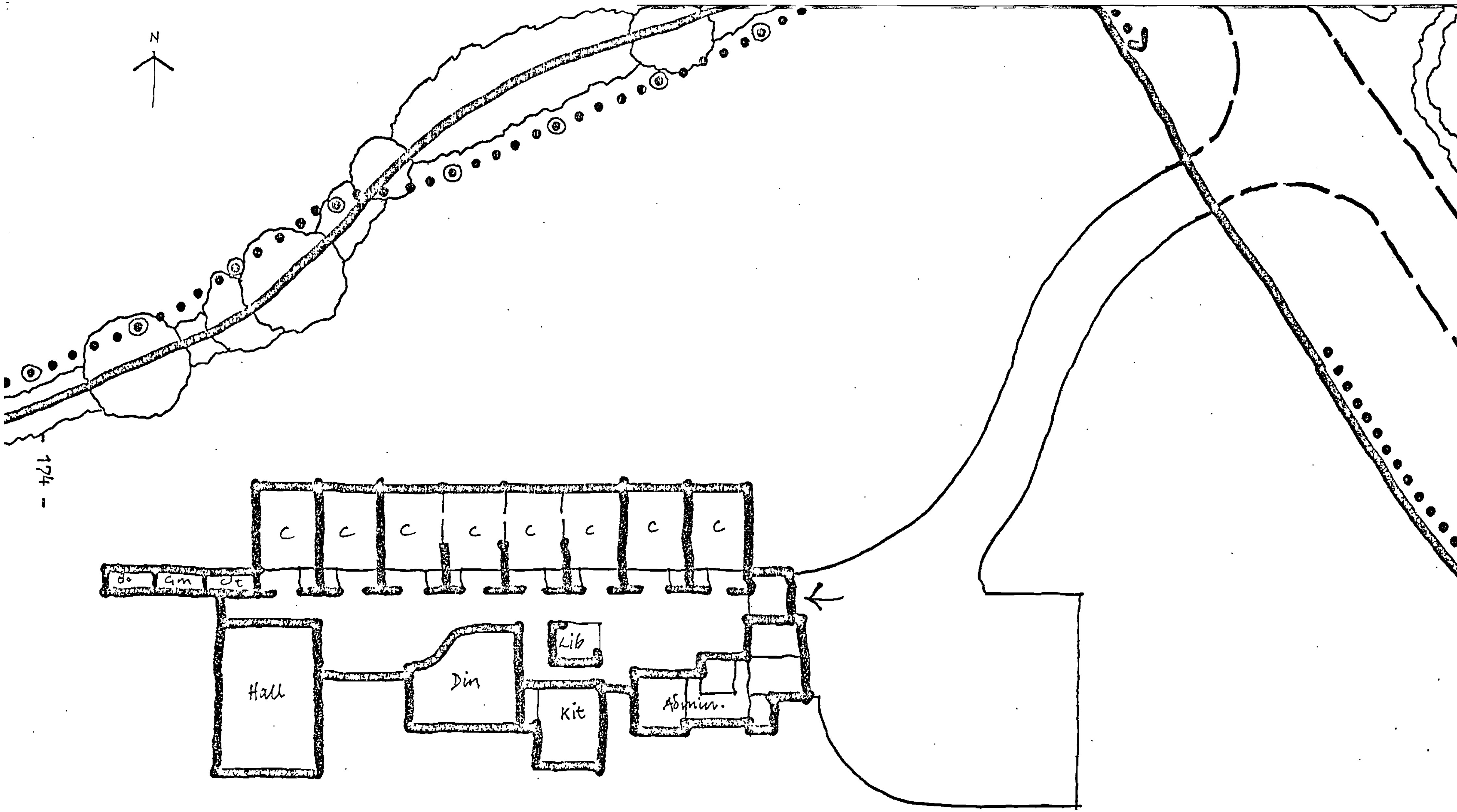
Key
+——— Overall evaluation eigenvector
×– – – Additive utility model index

Figure 8.2 The redrawn school plans

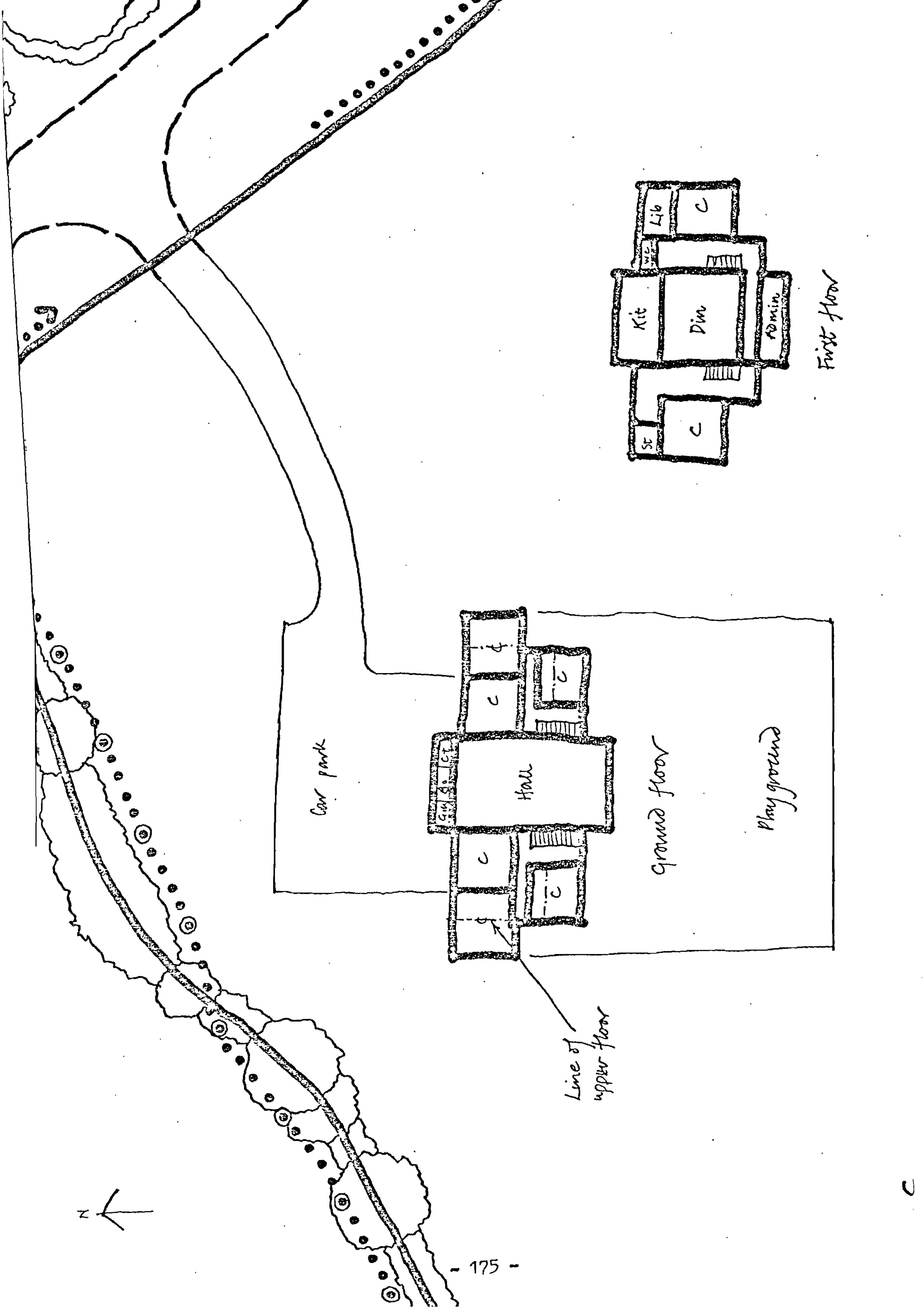
The following six pages show the redrawn school plans prepared in Experiment Four.

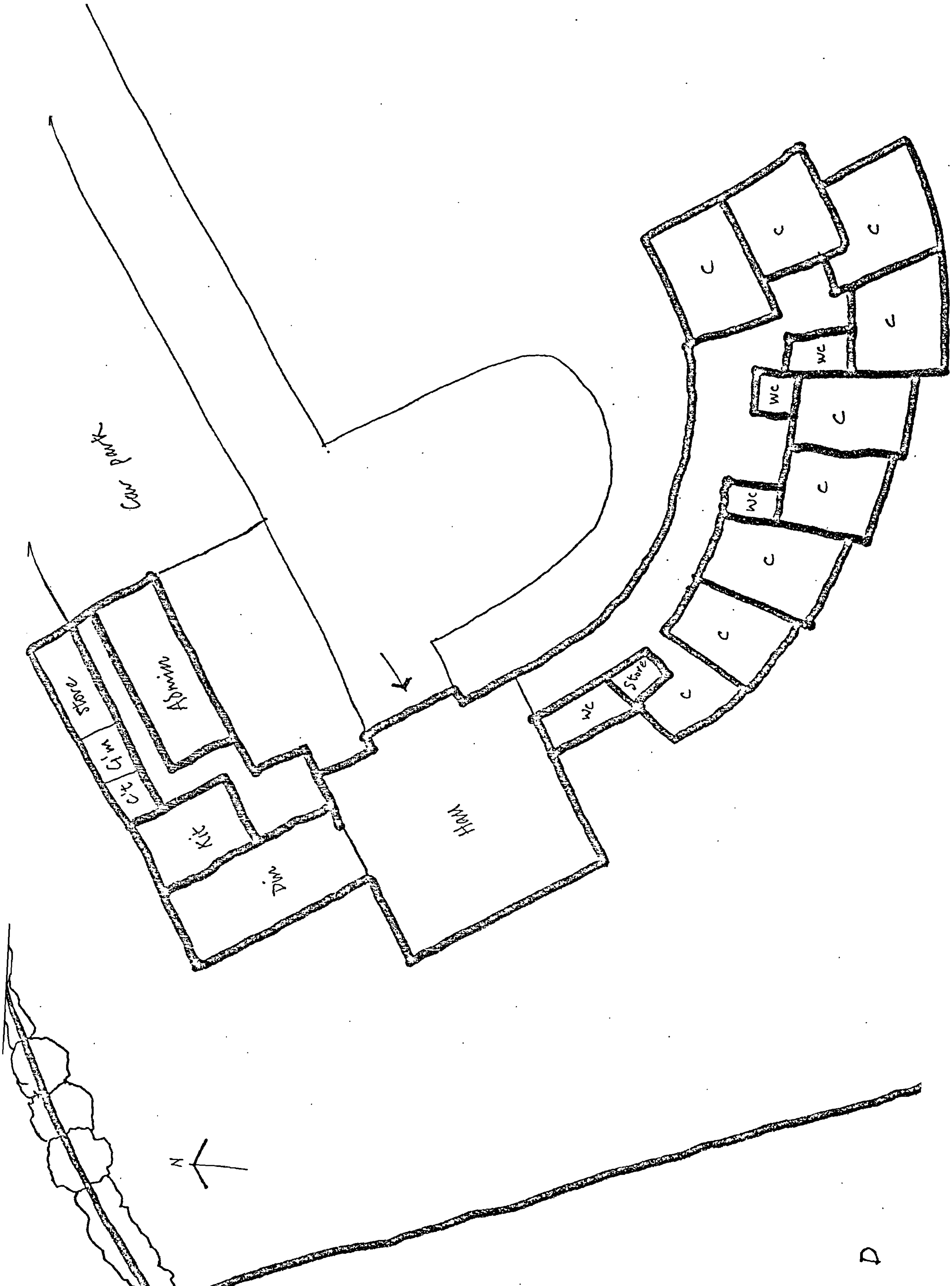


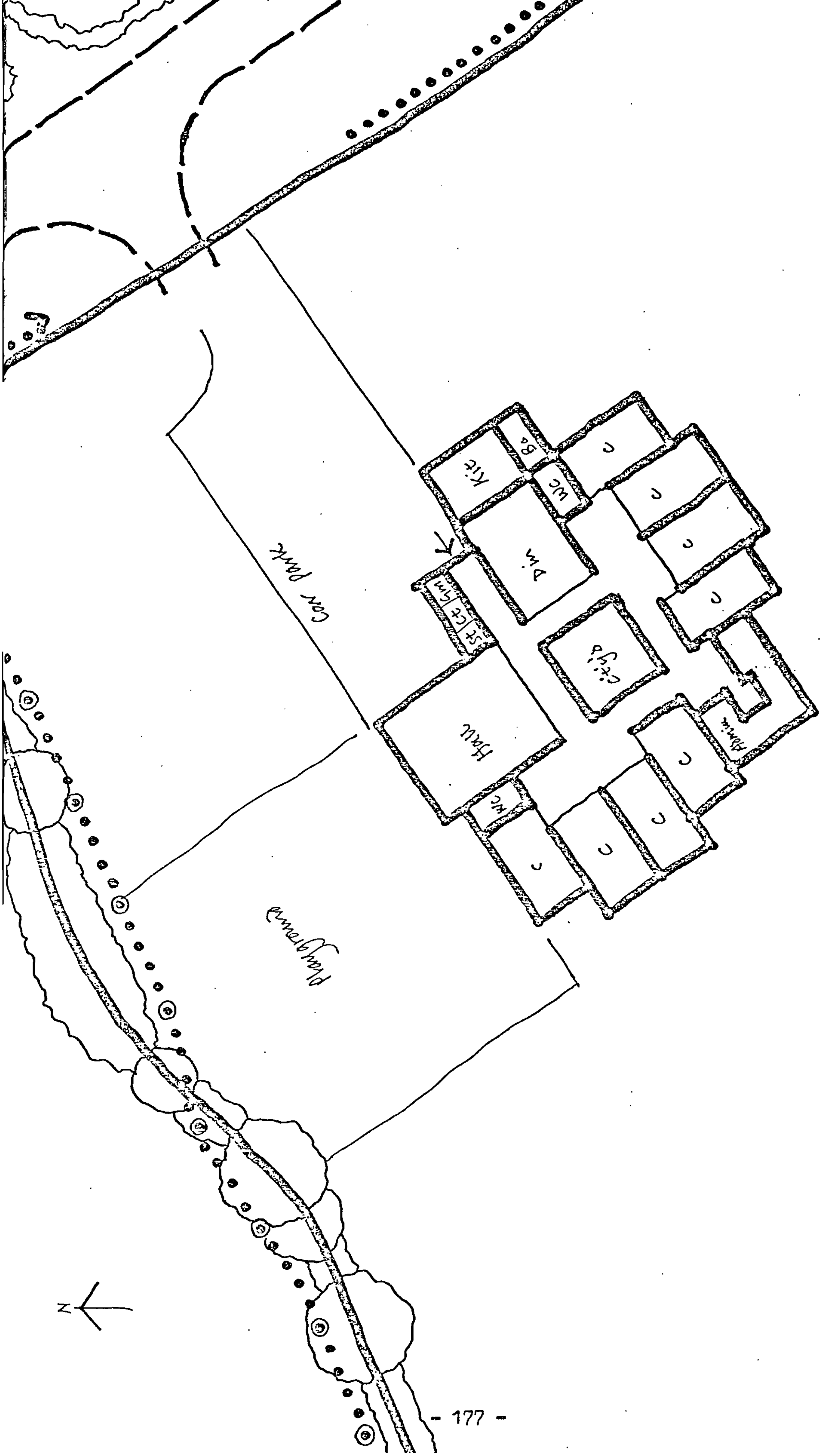
Car park
Admin
Din
Kit
Gym
HD
H
C
WC
Lib
Exhib
PT
Hall
N
A



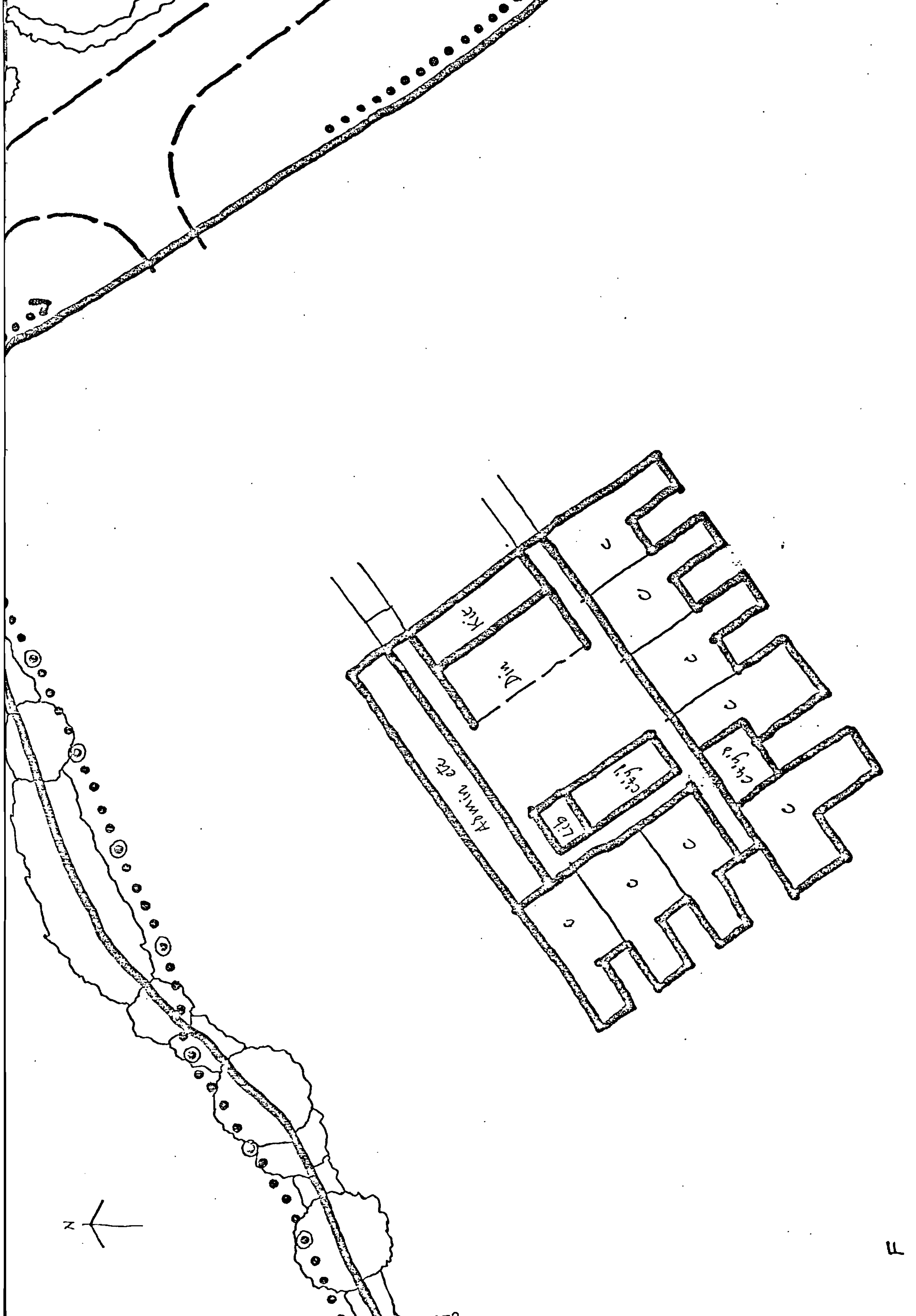
B


First floor
Kit
Din
Admin
Lib
WC
St
C
Car park
Hall
Ground floor
Playground
Line of upper floor
N


Car park
Store
Admin
Kit
Din
Hall
wc
Store
wc
wc
wc
C
C
C
C
C
C
C
C
C
N
D



E


Kit
Din
Admin etc
Lib
F

CHAPTER 9 EXPERIMENT FIVE

9.1 Introduction

The final experiment in the series uses the techniques that have been developed throughout the thesis. It also retains the basic form of experiment comprising the eliciting and scaling of attributes, an intensive design exercise, a repetition of the scaling of attributes after design, and concluding with the detailed subjective evaluation of alternative designs. However unlike Experiments Two, Three and Four, the six subjects are able to define their own subjective attributes independently. There is less concern with agreement in the choice of attributes and the level of concordance among the ratings of attributes, and more concern to find out what individual designers consider to be their priorities. So this exercise, while relying on the techniques which have so far been restricted to experimental settings with non-architects and students of architecture, is concerned with the possibilities for extending the use of these techniques towards the more realistic applications of identifying designers' priorities and of providing a means for formal and detailed subjective evaluations of alternative designs.

With these aims in mind the experiment has two important aspects. First all the subjects were qualified architects of between two and fifteen years experience. Second each subject was able to describe his own attributes both for design and for evaluation; he was not constrained that these should be the same, nor was he expected to confer and to agree attributes in a group situation.

Because each architect's own individual subjective attributes were elicited, comparisons between the subjects' priorities are not possible in the same way as for the previous experiments. Therefore many of the hypotheses which were tested in Experiments Three and Four are not tested here. Nevertheless some of the previous hypotheses are tested, including whether or not there is a difference between the internal consistency achieved in scaling attributes before and after design and whether or not the design process caused changes in the rating of priorities. The practice of using each subject as his own control by his scaling attributes for a holiday companion was retained. Between the subjects' evaluations two comparisons are made: whether there was significant concordance between their overall ratings of schemes, and whether there was significant concordance between the additive utility model indices.

9.2 Hypotheses

9.2.1 Correlations between the ratings of attributes before and after design

Hypothesis 1 That for each subject there would be no significant correlation between the rating of attributes for a holiday companion before design and the rating after design.

Hypothesis 2 That for each subject there would be no significant correlation between the rating of attributes for a school plan before design and the rating after design.

9.2.2 The effect of the order in which attributes were scaled

Hypothesis 3 That for each subject there would be no significant correlation between the rating of attributes for a holiday companion before design and the rating of attributes for a school plan before design, according to the order in which they were scaled.

Hypothesis 4 That for each designer there would be no significant correlation between the rating of attributes for a school plan after design and the rating of attributes for a holiday companion after design, according to the order in which they were scaled.

9.2.3 Differences between the consistency achieved in scaling attributes before and after design

Hypothesis 5 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a holiday comparion before design and that achieved in scaling attributes for a school plan before design.

Hypothesis 6 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a holiday companion before design and that achieved after design.

Hypothesis 7 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a school plan before design and that achieved after design.

Hypothesis 8 That there would be no significant difference between the set of consistency measures achieved in scaling attributes for a school plan after design and that achieved in scaling attributes for a holiday companion after design.

9.2.4 Concordance between evaluations of school plans

Hypothesis 9 That there would be no significant concordance between the subjects' rating of school plans given by the additive utility model indices.

Hypothesis 10 That there would be no significant concordance between the subjects' overall ratings of school plans.

9.2.5 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

Hypothesis 11 That for each subject there would be no significant correlation between the overall rating of school plans and the rating given by additive utility model indices.

Hypothesis 12 That for each subject there would be no correspondence between the overall rating of school plans and the hierarchical cluster analysis of similarity judgements of school plans.

Hypothesis 13 That for each subject there would be no correspondence between the additive utility model indices and the hierarchical cluster analysis of similarity judgements of school plans.

9.3 Experimental Method

In order to test the hypotheses the experiment was organised as follows:

1 The subjects (S5.1, S5.2, S5.3, S5.4, S5.5, S5.6) were interviewed individually by the experimenter. Each was given a brief verbal description of how the exercise would be conducted.

2 Each subject was read the following statement:

> "Consider the proposition of a fortnight's holiday in Greece. What important attributes or qualities would you look for in the selection of a holiday companion? Using brief descriptions write down the six most important ones onto the cards provided."

The cards were labelled 'a' to 'f'. The subject was then given the numerical scale to be used in rating attributes and a brief verbal description on how to scale paired comparisons. The experimenter then handed the subject all possible pairs of attributes in turn, in the order shown in appendix 4.2. A simple rank was also asked for and recorded.

3 Each subject was given the site plan and brief for a two-form entry primary school. He was read the following statement:

> "Consider the implications of planning a primary school, on the given site and to satisfy the given brief. What important attributes or qualities would you take into account in planning the school? Using brief descriptions write down the six most important ones onto the cards provided."

The cards were labelled 'a' to 'f'. Each subject was again given the numerical scale to be used in rating attributes. The experimenter then handed the subject all possible pairs of attributes in turn, and recorded the scaled judgements for each pair using the form shown in appendix 4.2. A simple rank order was also asked for and recorded.

4 Each subject then proceeded to design a school plan on the given site and to satisfy the given brief. The subjects took between two and four hours, with an average of about three hours.

5 Each subject was asked again about the attributes he felt to be important in the planning of a school and was given the option of amending those he had offered before design. All six subjects elected to retain the attributes offered before design. Each subject scaled paired comparisons of his own attributes using the same method as before design, as described above.

6 Each subject scaled paired comparisons of his own attributes for a holiday companion using the same method as before design, as described above.

7 The six school plans were redrawn by the experimenter to a standard scale (1:500), orientation and format (figure 9.2).

8 The redrawn plans were evaluated by the six subjects. Each worked individually.

Each subject was shown all six plans briefly. Then the form (appendix 4.6) with instructions on how to make similarity judgements of triadic comparisons of school plans was given to the subject. The experimenter worked through and gave the subject all possible triads of school plans in the order shown on the form. In each case the subject was given the choice of completing the form himself or having the experimenter do so. Subjects S5.2, S5.4 and S5.6 themselves recorded for each triad the similar pair and the attribute used to separate this pair from the third; subjects S5.1, S5.3 and S5.5 asked the experimenter to do so. It was not considered significant which course of action was followed.

9 The attributes used to differentiate between plans were the attributes used in the evaluation of the plans. Each subject copied his own attributes for evaluation onto cards, and labelled them alphabetically. Each subject was then asked to scale these attributes in terms of their relative importance in the planning of a school. Each subject was given the numerical scale to be used. The experimenter then handed the subject all possible pairs of attributes in turn, and recorded the scaled judgement for each pair. A simple rank order was also recorded.

10 The redrawn plans were evaluated by each subject with respect to each of his own attributes for evaluation. Each subject was given the numerical scale to be used. Taking each attribute in turn, the experimenter handed all possible pairs of school plans to the subject and recorded the scaled judgement for each pair with respect to that attribute. In this way all the attributes were worked though. The attributes were taken in the order in which they had been lettered alphabetically after being elicited in the triadic comparisons. This was not necessarily the order of their subjective importance. A simple rank order of plans with respect to each attribute was also recorded.

11 Finally the redrawn plans were evaluated by each subject for their overall merit as schools. Each subject was given the numerical scale to be used. The experimenter handed all possible pairs of school plans to the subject in turn, in the order shown in appendix 4.4, and recorded the scaled judgement for each pair.

9.4 Results

9.4.1 Correlation between the ratings of attributes before and after design

Hypothesis 1 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for a holiday companion before and after design. Table 9.3 shows the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level.

Subject	Correlation coefficient	Sig. level
S5.1	0.986	.01
S5.2	0.900	.05
S5.3	1.000	.01
S5.4	0.600	NS
S5.5	0.743	NS
S5.6	0.943	.01

In the cases of four of the six subjects, S5.1, S5.2, S5.3, S5.6, there was significant correlation. The other two subjects, S5.4 and S5.5, achieved a positive correlation coefficient although in neither case was it statistically significant. However it is of interest that both these two subjects exhibited low levels of internal consistency in scaling the paired comparisons, as shown by the consistency measures in table 9.5. A comparison between the simple ranks before and after design for these two subjects gives the following:

Subject	Correlation coefficient	Sig. level
S5.4	0.886	.05
S5.5	0.943	.01

The null hypothesis was rejected.

Hypothesis 2 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for school plan design before and after design. Table 9.4 shows the data. The following tabulation shows for each subject the rank correlation coefficient and the significance level.

Subject	Correlation coefficient	Sig. level
S5.1	0.829	.05
S5.2	0.829	.05
S5.3	1.000	.01
S5.4	1.000	.01
S5.5	1.000	.01
S5.6	0.829	.05

The null hypothesis was rejected.

9.4.2 The effect of the order in which attributes were scaled

Hypothesis 3 was tested by calculating for each subject Spearman's correlation coefficient between the ranks of attributes for a holiday companion before design and the ranks of attributes for school plan design before design according to the order in which they were scaled. Tables 9.3 and 9.4 show the data. The following tabulation shows for each subject the rank correlation coefficient, and the signficiance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S5.1	0.814	NS
S5.2	0.429	NS
S5.3	0.314	NS
S5.4	-0.371	NS
S5.5	-0.371	NS
S5.6	-0.086	NS

The null hypothesis was not rejected.

Hypothesis 4 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of attributes for school plan design after design and the ranks of attributes for a holiday companion after design according to the order in which they were scaled. Tables 9.3 and 9.4 show the data. The following tabulation shows for each subject the rank correlation coefficient, and the significance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S5.1	0.486	NS
S5.2	0.671	NS
S5.3	0.314	NS
S5.4	0.086	NS
S5.5	-0.600	NS
S5.6	-0.257	NS

The null hypothesis was not rejected.

9.4.3 Differences between the consistency achieved in scaling attributes before and after design

Hypotheses 5 to 8 were tested by calculating Mann-Whitney's U between the pairs of sets of consistency measures being compared. Table 9.5 shows the data. The following tabulation shows for each hypothesis (H) the pairs of sets of eigenvalues being compared, the corresponding value of U, and the probability under the null hypothesis (one-tailed test).

H	Before design		After design		U	Probability under Ho
	Holiday companion	School plan	School plan	Holiday companion		
5	x	x			11	0.155
6	x			x	10	0.120
7		x	x		12.5	0.220
8			x	x	14	0.294

Null hypothesis 5 was not rejected.
Null hypothesis 6 was not rejected.

Null hypothesis 7 was not rejected.
Null hypothesis 8 was not rejected.

9.4.4 Concordance between evaluations of school plans

Hypothesis 9 was tested by calculating Kendall's coefficient of concordance between the six sets of ranks of the subjects' rating of school plans given by additive utility model indices. Table 9.8 shows the data. The coefficient was found to be 0.149 which was not significant. The null hypothesis was not rejected.

Hypothesis 10 was tested by calculating Kendall's coefficient of concordance between the six sets of ranks of the subjects' overall ratings of school plans. Table 9.9 shows the data. The coefficient was found to be 0.251 which was not significant. The null hypothesis was not rejected.

9.4.5 Overall rating of school plans, additive utility model indices and hierarchical cluster analysis

Hypothesis 11 was tested by calculating for each subject Spearman's rank correlation coefficient between the ranks of overall ratings of school plans and the ranks of additive utility model indicies. The following tabulation shows for each subject the rank correlation coefficient and the significance level (one-tailed test).

Subject	Correlation coefficient	Sig. level
S5.1	0.829	.05
S5.2	0.600	NS
S5.3	0.600	NS
S5.4	0.771	NS
S5.5	0.943	.01
S5.6	0.943	.01

In all six cases there was a positive correlation coefficient but for only three of the subjects, S5.1, S5.5 and S5.6, was it significant. The null hypothesis could not be entirely rejected.

Hypotheses 12 and 13 were tested by comparing for each subject the hierarchical cluster analysis of similarity judgements with both the overall ratings of school plans and the additive utility model indices. Figure 9.1 shows the data. The following tabulation gives a verbal extimate, on a four point scale, 'high', 'medium', 'low', 'no', of the correspondence for each subject between the hierarchical cluster analysis, the overall rating and the indices.

Subject	Correspondence with hierarchical analysis	
	Overall rating	Additive utility model indices
S5.1	High correspondence. Cluster A-B-E: all rated highly. Cluster C-F: both rated low.	Medium correspondence. Cluster A-B-E: all rated highly. Cluster C-F: both rated low.
S5.2	High correspondence. Cluster A-B: both rated highly. Cluster C-D-E-F: all rated low.	High correspondence. Cluster A-B: both rated highly. Cluster C-D-E-F: all rated low.
S5.3	High correspondence. Cluster B-C: both rated highly. E,A: rated third, fourth. Cluster D-F: both rated low.	Low correspondence. Cluster B-C: both rated highly.
S5.4	Medium correspondence. Cluster A-B-D-F: all rated highly.	Medium correspondence. Cluster A-B-D-F: all rated highly.
S5.5	No correspondence.	Low correspondence. Cluster C-D: both rated similarly. Cluster A-B: both rated similarly.
S5.6	Medium correspondence. Cluster D-F: both rated highly. Cluster B-E: both rated low. C rated near B-E.	High correspondence. Cluster D-F: both rated highly. A rated near D-F. Cluster B-E: both rated low. C rated near B-E.

Null hypothesis 12 was rejected.

Null hypothesis 13 was rejected.

9.5 Conclusions

9.5.1 The effect of the design process on rating attributes

The purpose of each subject acting as his own control by scaling attributes for a holiday companion was to ensure that if changes were found in the rating of attributes for a school plan before and after design, this change could definitely be ascribed to the design process, assuming that the rating of attributes for a holiday companion had not changed. The conclusion to be drawn from the results of testing hypothesis 1 is that for each subject there was no significant change in the rating of attributes for a holiday companion. The conclusion to be drawn from the results of testing hypothesis 2 is that for each subject there was no significant change in the rating of attributes caused by the design process.

9.5.2 The effect of the order in which the attributes were scaled

The results of testing hypotheses 3 and 4 were as expected. They show that the order in which the pairs of attributes were scaled by the subjects had no effect on the scaled judgements and therefore on the ratings of attributes.

9.5.3 The effect of the design process on consistency

The results of testing hypotheses 5, 6, 7 and 8 were as expected in view of the findings of Experiments Three and Four. There was no significant difference between the set of consistency measures achieved in scaling attributes for a holiday companion before design compared with those achieved in scaling attributes for a school plan before design. There was no significant difference between the set of consistency measure achieved in scaling attributes for a holiday companion before design compared with after design. There was no significant difference between the set of consistency measures achieved in scaling attributes for a school plan after design compared with those achieved in scaling attributes for a holiday companion after design. There was no significant difference between the set of consistency measures achieved in scaling attributes for a school plan before design compared with after design. The process of designing a school plan did not cause changes in the consistency with which attributes for school plan design were scaled.

9.5.4 Concordance between evaluations of school plans

There was no significant concordance between the subjects' evaluations of school plans, either in the ratings of overall merit or the ratings given by the additive utility model indices. As there was considerable variety in the attributes elicited from the subjects, and as the six architects may be presumed to have high levels of design skill and therefore to be able to express accurately their priorities in their designs, the lack of significant concordance between their evaluations is the expected result.

9.5.5 Overall ratings of school plans, additive utility model indices and hierarchical cluster analysis

The conclusions which may be drawn from the results of testing hypotheses 11, 12 and 13 concern the comparison between three methods of evaluation of the same stimulus items, the school plans. The following tabulation is a combination of the results of testing hypotheses 11, 12 and 13. For each of the three comparisons between the overall merit ratings, the additive utility model indices and the hierarchical cluster analysis, each subject is given a verbal extimate on a four point scale, 'high', 'medium', 'low', 'no', of the degree of correspondence. In this tabulation the rank correlation coefficients found in testing hypothesis 11 are converted to the verbal description thus: correlation significant at 0.01 level, 'high'; correlation significant at 0.05 level, 'medium'; positive correlation, 'low'; negative correlation, 'no'.

	Degree of correspondence between pairs of results		
Overall rating	x	x	
Indices	x		x
Cluster analysis		x	x
S5.1	medium	high	medium
S5.2	low	high	high
S5.3	low	high	low
S5.4	low	medium	medium
S5.5	high	no	low
S5.6	high	medium	high

In this table it should be remembered that it takes only one set of results to fail to correspond with either of the other two sets for this to affect two of the three pairs. The results are therefore encouraging in that in all cases except S5.4 there is at least one pair of results that shows a high degree of correspondence. Because in all cases except S5.4 there is only one set of results which does not correspond to a high degree with the other two, this set can be identified for each subject thus:

S5.2 Utility model indices failed to correspond

S5.3 Utility model indices failed to correspond

S5.5 Cluster analysis failed to correspond

S5.4's results exhibited the least correspondence among the six subjects' sets of results. Part of the discrepancy may be ascribed to his rating of attributes for evaluation; although there were just six of these, the consistency measure was 7.437, signifying serious inconsistencies in the rating of attributes and leading to the lack of correspondence between the overall merit ratings of school plans and the ratings given by additive utility model indices. Even in the overall merit rating of school plans consistency was below average at 6.774.

S5.6's results exhibited the most correspondence, though there was a slight discrepancy between his overall rating of plans and the similarity judgements.

These results are highly encouraging. They imply that the techniques used have the potential to provide means of eliciting, structuring and representing subjective judgements. Although certain discrepancies have been detected, there is little doubt, particularly on the evidence shown in figure 9.1, that the additive utility model indices do express preferences among design alternatives within certain tolerances of accuracy, and that in some cases these tolerances are very fine. The accuracy of the indices suggests that their method of computation based on the weighting of attributes and the weighting of plans with respect to attributes presents a numerically precise and logically coherent set of principles for the subjective evaluation of designs.

Table 9.1

Description of attributes for a holiday companion

Subject	Attribute	Description
S5.1	a	Sense of humour
	b	Similarity of interests (lack of dissimilarity)
	c	Similar financial priorities
	d	Non-smoker
	e	Optimistic outlook
	f	Friendly disposition to others
S5.2	a	Easy going
	b	Feel reliant financially
	c	Adventurous mind/approach
	d	Interest in people's indigenous lifestyles
	e	Sharing of enjoyment/activities
	f	Liking of sun, sea, food
S5.3	a	Availability - can they go away when I want to
	b	Compatability - do they enjoy doing similar things to me, e.g. sunbathing or visiting places
	c	Money - do they have a similar amount to spend
	d	Initiative - can they take decisions on they own
	e	Humour - is theirs the same as mine
	f	Adaptability - are they prepared to rough it as well as stay in hotels
S5.4	a	Wide interests
	b	Female
	c	Energetic
	d	Enthusiastic
	e	Considerate
	f	Organised
S5.5	a	Self reliance
	b	Common sense of humour
	c	Some common interests
	d	About same financial limit
	e	Similar sense of adventure
	f	Mutually considerate - sense of conpromise

Sub-ject	Attri-bute	Description
S5.6	a	Openmindedness - critical and enquiring attitude to new experiences
	b	Anti-garrulousness
	c	Likes food and new sorts
	d	Not defensive towards the nations
	e	Knowledgeable about places visited
	f	Ability to relax and be lazy

Table 9.2

Description of attributes for school plan design

Subject	Attribute	Description
S5.1	a	Assembly hall on main circulation route, also direct access to playground
	b	Classroom - relation to outside environment including courtyards (using outside as teaching area), (views, access and sunshine)
	c	Dining areas - linked to kitchen, stores and delivery area
	d	Administration areas - some seclusion from noisy child areas
	e	Library - a common resource; therefore on main circulation; some quieter area
	f	Boiler - type of heating system? running costs/energy
S5.2	a	Flexibility - short term: daily use/activities - mid term: changing educational approaches - long term: possible other uses (non school?)
	b	Optimise use of scarce resources - each space to offer alternative possible uses - contiguous spaces jointly offering other alternatives - waste eliminated - internal circulation - external vehicular access
	c	Child's scale/identification - recognisable/different spaces - clarity of space organisation - group identity
	d	Outside/inside relationships - teaching extends outside - openness
	e	Orientation/aspect/shadowing
	f	Access - vehicles: car park (staff, visitors), deliveries - pedestrians
S5.3	a	Orientation - relating the building to the sun, prevailing winds and views
	b	Front entrance - making it clear to a visitor how/where to enter the building
	c	Flexibility - potential for re-arranging class sizes and carrying out internal replanning
	d	Money - its effect on plan, form and circulation areas
	e	Informality/scale - helping to create an environment in which studying is encouraged, by planning
	f	Grouping - linking rooms/areas of related activities near to each other

Subject	Attribute	Description
S5.4	a	Closely articulated areas - densly planned core to the school
	b	Secure 'family' spaces - each child has an opportunity of identifying his class/space - internal and external areas
	c	Indication of hierarchy - headmistress and teachers at the centre of the school - clarity of control
	d	Choice of private spaces or loose open meeting/circulation areas i.e. an adaptable building
	e	Flexibility in the classroom - anopportunity for all to structure/change/rebuild their own environment for different activities, i.e. Adventure
	f	Formal assembly hall - visual discipline of ranks/grades within the school. Unchanging focus of the essence of the school
S5.5	a	Economy of circulation: reducing and making optimum use (compactness)
	b	Grouping of spaces by function, time and frequency of use and by whom
	c	Provision of sunlight, priority: classrooms, library, administration, dinning, assembly
	d	Separate service/personnel entrance
	e	Relationship to outside for views and daylight, priority: classrooms, library, dining, assembly, administration
	f	Breaking down overall massing to create outdoor space and to avoid over-simple form
S5.6	a	Similarity of parts
	b	Administration hierarchy <u>not</u> reflected in plan arrangement
	c	Flexibility
	d	Light and airy (as opposed to 'cosy')
	e	Simple structure/building system
	f	Clear, easily understandable circulation

Table 9.3 Rating of attributes for a holiday companion before and after design: eigenvectors, corresponding ranks and rank correlation coreeicient (r)

Subject	Attribute	Before design		After design		Correlation	
		Eigen-vector	Rank	Eigen-vector	Rank	r	Significance level
S5.1	a b c d e f	.111 .385 .228 .111 .099 .066	3.5 1 2 3.5 5 6	.131 .337 .201 .162 .118 .052	4 1 2 3 5 6	0.986	0.01
S5.2	a b c d e f	.074 .369 ,298 .112 .074 .074	5 1 2 3 5 5	.071 .357 .357 .071 .071 .071	4.5 1.5 1.5 4.5 4.5 4.5	0.900	0.05
S5.3	a b c d e f	.381 .230 .101 .032 .184 .073	1 2 4 6 3 5	.376 .221 .105 .042 .191 .066	1 2 4 6 3 5	1.000	0.01
S5.4	a b c d e f	.364 .059 .120 .189 .232 .037	1 5 4 3 2 6	.136 .063 .177 .222 .370 .032	4 5 3 2 1 6	0.600	NS
S5.5	a b c d e f	.234 .093 .067 .224 .092 .291	2 4 6 3 5 1	.237 .069 .071 .160 .075 .388	2 6 5 3 4 1	0.743	NS
S5.6	a b c d e f	.465 .202 .115 .047 .033 .139	1 2 4 5 6 3	.429 .125 .089 .043 .034 .279	1 3 4 5 6 2	0.943	0.01

Table 9.4 Rating of attributes for school plan design before and after design: eigenvectors, corresponding ranks and rank correlation coefficient (r)

Subject	Attribute	Before design		After design		Correlation	
		Eigen-vector	Rank	Eigen-vector	Rank	r	Significance level
S5.1	a b c d e f	.216 .514 .108 .063 .070 .030	2 1 3 5 4 6	.240 .434 .075 .080 .141 .030	2 1 5 4 3 6	0.829	0.05
S5.2	a b c d e f	.181 .172 .431 .066 .057 .093	2 3 1 5 6 4	.193 .166 .485 .056 .052 .048	2 3 1 4 5 6	0.829	0.05
S5.3	a b c d e f	.427 .150 .070 .102 .055 .196	1 3 5 4 6 2	.346 .177 .061 .105 .046 .266	1 3 5 4 6 2	1.000	0.01
S5.4	a b c d e f	.037 .339 .051 .279 .212 .083	6 1 5 2 3 4	.034 .349 .050 .272 .241 .054	6 1 5 2 3 4	1.000	0.01
S5.5	a b c d e f	.120 .376 .057 .160 .249 .037	4 1 5 3 2 6	.077 .381 .049 .187 .260 .046	4 1 5 3 2 6	1.000	0.01
S5.6	a b c d e f	.055 .088 .260 .034 .228 .335	5 4 2 6 3 1	.056 .076 .237 .033 .314 .285	5 4 3 6 1 2	0.829	0.05

Table 9.5 Consistency achieved in scaling attributes

Subject	Before design				After design			
	Holiday companion		School plan		School plan		Holiday companion	
	λmax	Rank	λmax	Rank	λmax	Rank	λmax	Rank
S5.1	6.943	4	6.411	3	6.546	5	6.725	4
S5.2	6.320	1	6.329	2	6.158	1	6.000	1
S5.3	6.582	2	6.202	1	6.345	3	6.374	3
S5.4	7.389	6	6.735	4	6.329	2	7.078	6
S5.5	7.313	5	7.124	6	6.709	6	6.277	2
S5.6	6.739	3	6.812	5	6.843	4	6.756	5

Table 9.6 Description of attributes for school plan evaluation

Sub-ject	Attri-bute	Eigen-vector	Description
S5.1	a	.192	Circulation - efficiency in moving - shape positive/negative - aesthetic
	b	.242	Classrooms - aspect/orientation
	c	.032	Handling of service areas, kitchen waste/fuel
	d	.102	Courtyard - effectiveness
	e	.066	Consideration of staff rooms - location, off circulation
	f	.197	Architectural content - giving the building a focus - creation internal places - connections with the site
	g	.045	Entering the building - children, staff
	h	.042	Library - location/handling
	i	.081	Building economies
	Consistency: $\lambda max = 9.603$		
S5.2	a	.189	Cost
	b	.024	External access
	c	.055	Internal circulation
	d	.270	Child identification - group, scale
	e	.119	Flexibility
	f	.025	Orientation
	g	.249	Richness of spatial relations/experience, including inside/outside relations, broken up geometry
	h	.069	Compactness - energy
	Consistency: $\lambda max = 8.906$		
S5.3	a	.047	Compactness
	b	.142	Outward growing discipline - planned informality
	c	.209	Clear main entrance to building
	d	.096	Clear circulation route
	e	.153	Relation between plan, form and site
	f	.288	Focal point or central space
	g	.064	Zoning of activities
	Consistency: $\lambda max = 7.496$		

Subject	Attribute	Eigen vector	Description
S5.4	a	.158	Good orientation of elements
	b	.317	Good integration and grouping of elements
	c	.228	Variety, interest, choice and informality of layout
	d	.083	Symmetry, order, simplicity and formality of layout
	e	.093	Cheap to construct, compact
	f	.121	Good circulation
	Consistency: $\lambda max = 7.473$		
S5.5	a	.154	Exploitation of external spaces - use
	b	.097	Main entrance - staff overlooking and access to staff
	c	.487	Internal circulation - separate from hall
	d	.043	Exploit internal space - relationship to courtyards, visual and circulation
	e	.049	Compactness
	f	.170	Views and sunlight - classroom, hall, staff
	Consistency: $\lambda max = 6.817$		
S5.6	a	.449	Orientation
	b	.218	Compactness
	c	.068	Symmetry
	d	.265	Clear geometry
	Consistency: $\lambda max = 4.282$		

Table 9.7 Rating of school plans with respect to attributes, and consistency achieved

Subject	Attribute	Plans Eigenvectors A	B	C	D	E	F	Consistency λmax
S5.1	a	.370	.139	.096	.073	.277	.046	6.439
	b	.243	.271	.051	.089	.249	.098	6.334
	c	.354	.129	.068	.107	.226	.116	6.350
	d	.331	.303	.060	.075	.185	.046	6.322
	e	.392	.109	.162	.056	.083	.197	6.277
	f	.374	.203	.068	.115	.199	.040	6.417
	g	.317	.177	.087	.198	.172	.048	6.690
	h	.300	.144	.051	.145	.289	.071	6.233
	i	.050	.130	.232	.077	.141	.369	6.328
S5.2	a	.031	.194	.251	.076	.057	.391	6.663
	b	.201	.488	.038	.125	.098	.050	6.613
	c	.243	.486	.050	.098	.059	.064	6.492
	d	.356	.356	.035	.084	.136	.033	6.434
	e	.040	.295	.103	.107	.046	.410	6.461
	f	.067	.290	.029	.419	.073	.122	6.695
	g	.429	.296	.045	.119	.082	.028	6.623
	h	.030	.245	.149	.110	.057	.409	6.442
S5.3	a	.032	.108	.248	.159	.058	.395	6.484
	b	.112	.429	.218	.042	.168	.032	6.433
	c	.036	.177	.247	.397	.085	.058	6.501
	d	.031	.104	.366	.159	.055	.286	6.677
	e	.049	.255	.448	.052	.150	.047	6.234
	f	.280	.095	.377	.162	.046	.039	6.236
	g	.139	.034	.258	.104	.053	.412	6.577
S5.4	a	.257	.127	.034	.268	.257	.057	6.278
	b	.171	.194	.030	.469	.060	.076	6.642
	c	.251	.343	.049	.211	.102	.043	7.023
	d	.144	.129	.032	.402	.067	.225	6.842
	e	.031	.138	.096	.200	.059	.476	6.640
	f	.263	.377	.041	.204	.052	.053	6.783
S5.5	a	.460	.267	.040	.046	.137	.050	6.297
	b	.037	.215	.298	.110	.297	.043	6.424
	c	.092	.041	.179	.220	.419	.049	6.450
	d	.292	.145	.062	.093	.362	.045	6.323
	e	.035	.060	.223	.186	.099	.397	6.476
	f	.131	.313	.058	.078	.290	.130	6.593
S5.6	a	.184	.084	.025	.279	.133	.296	6.342
	b	.033	.067	.282	.177	.097	.345	6.644
	c	.186	.055	.032	.343	.038	.345	6.358
	d	.108	.037	.184	.296	.038	.337	6.352

Table 9.8 Rating of school plans given by additive utility model indices, corresponding ranks and coefficient of concordance (W)

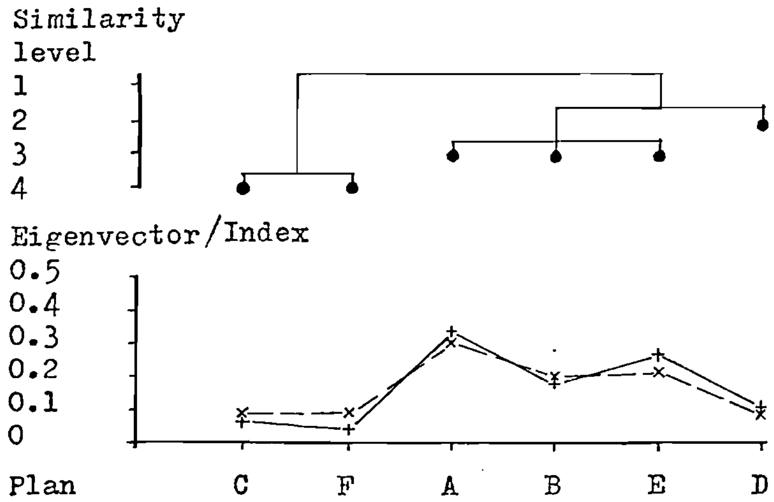
Plan	Subjects												Concordance	
	Indices						Ranks						W	Sig. level
	S5.1	S5.2	S5.3	S5.4	S5.5	S5.6	S5.1	S5.2	S5.3	S5.4	S5.5	S5.6		
A	.306	.236	.125	.199	.156	.131	1	2	4	3	2	3		
B	.199	.304	.182	.229	.144	.066	3	1	2	2	5	6		
C	.088	.095	.323	.043	.146	.124	6	5	1	6	4	4	.149	NS
D	.094	.106	.173	.316	.151	.266	5	4	3	1	3	2		
E	.216	.085	.089	.100	.324	.094	2	6	6	5	1	5		
F	.097	.175	.108	.114	.079	.321	4	3	5	4	6	1		

Table 9.9 Overall rating of school plans: eigenvectors, corresponding ranks and coefficient of concordance (W)

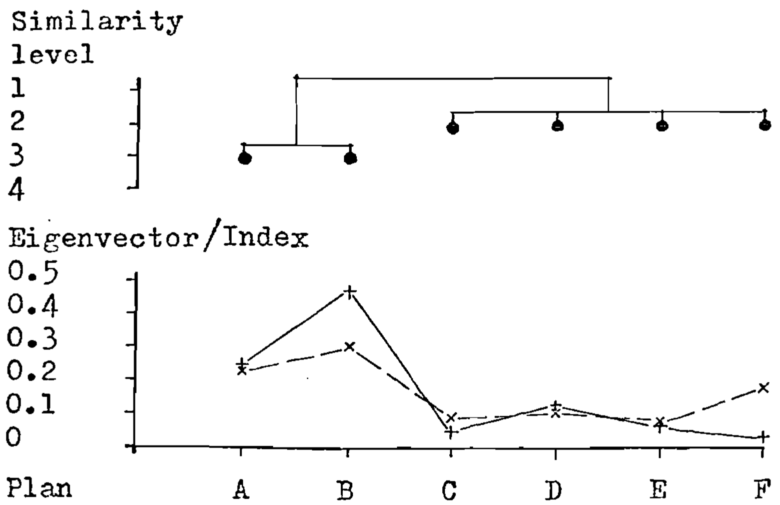
Plan	Subjects												Concordance	
	Eigenvectors						Ranks							
	S5.1	S5.2	S5.3	S5.4	S5.5	S5.6	S5.1	S5.2	S5.3	S5.4	S5.5	S5.6	W	Sig. level
A	.330	.251	.068	.394	.266	.078	1	2	4	1	2	4		
B	.180	.476	.247	.152	.103	.033	3	1	2	3	4	6		
C	.068	.045	.478	.029	.070	.147	5	5	1	6	5	3	.251	NS
D	.106	.127	.056	.319	.129	.227	4	3	5	2	3	2		
E	.271	.071	.120	.062	.395	.050	2	4	3	4	1	5		
F	.045	.031	.030	.045	.037	.466	6	6	6	5	6	1		

Figure 9.1 Comparison between each subject's similarity judgments of school plans and the weighting given in overall evaluation and in evaluation derived from additive utility model indices.

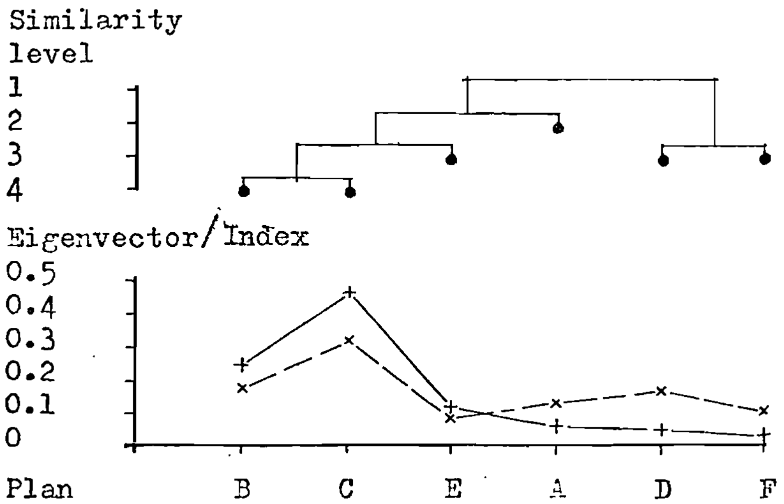
Subject S5.1



Subject S5.2



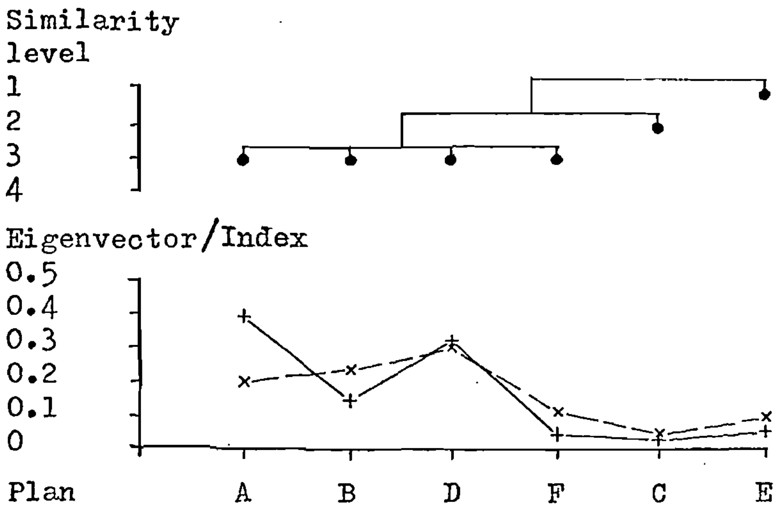
Subject S5.3



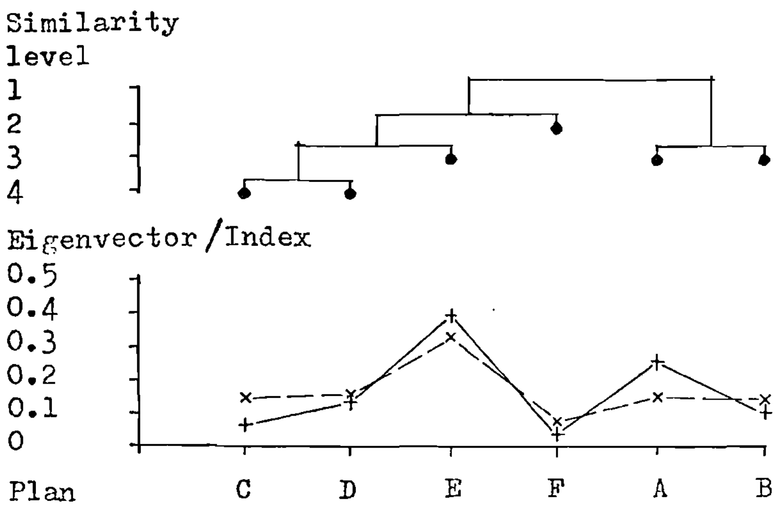
Continued

Figure 9.1 continued

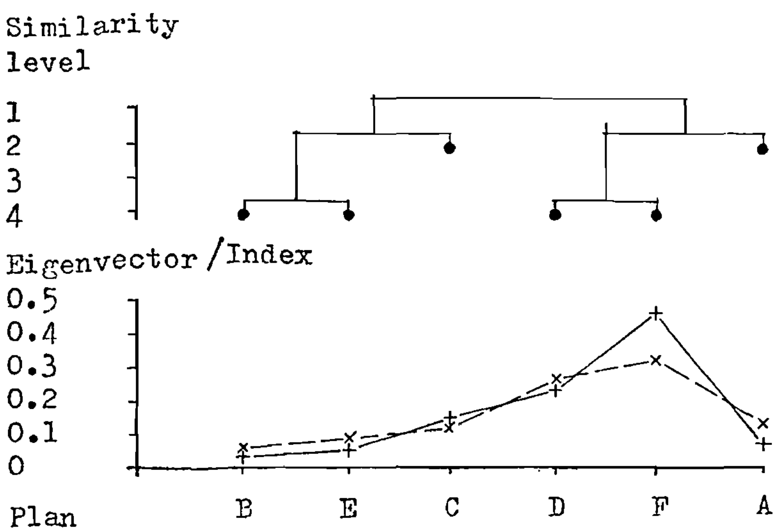
Subject S5.4



Subject S5.5



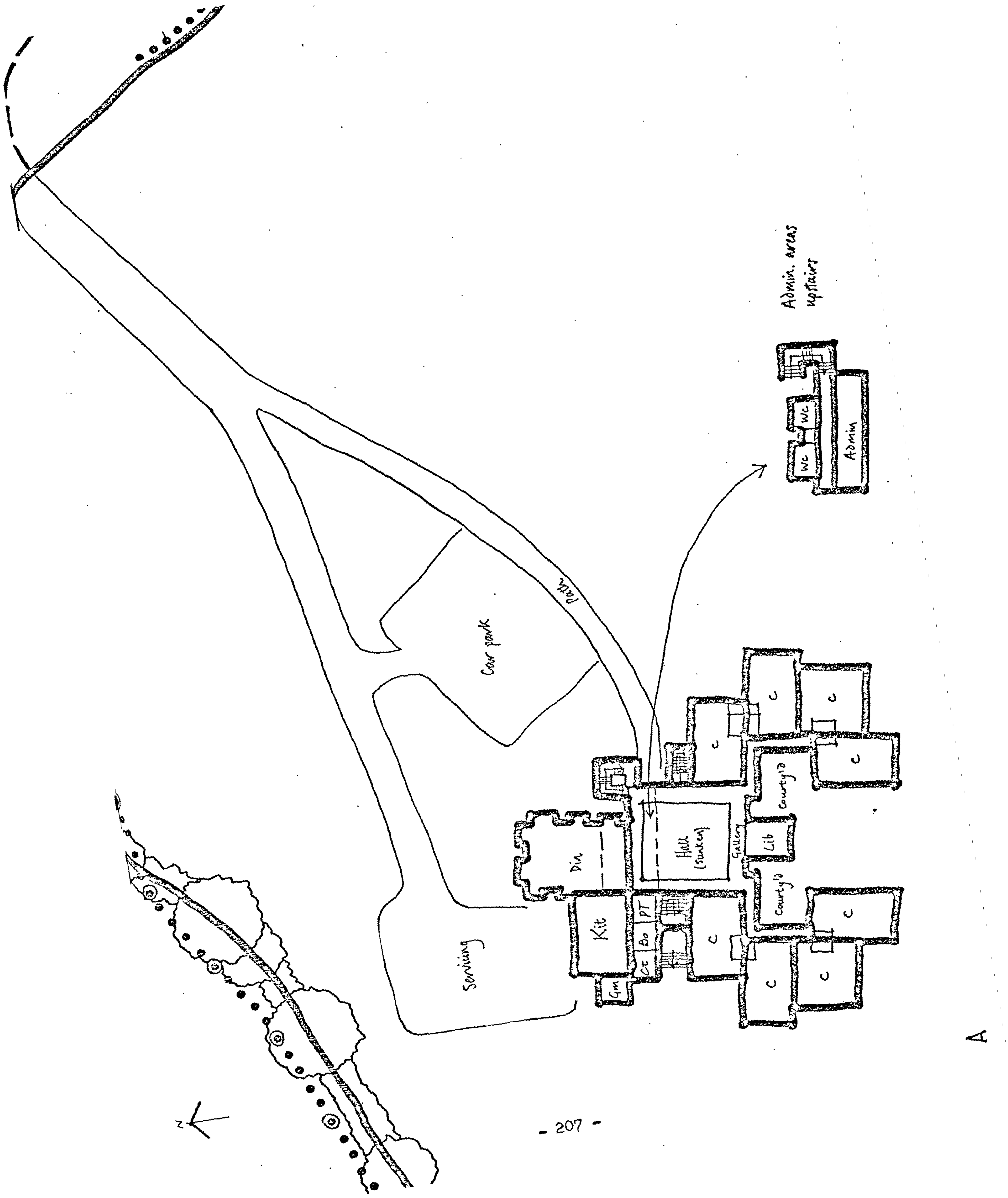
Subject S5.6

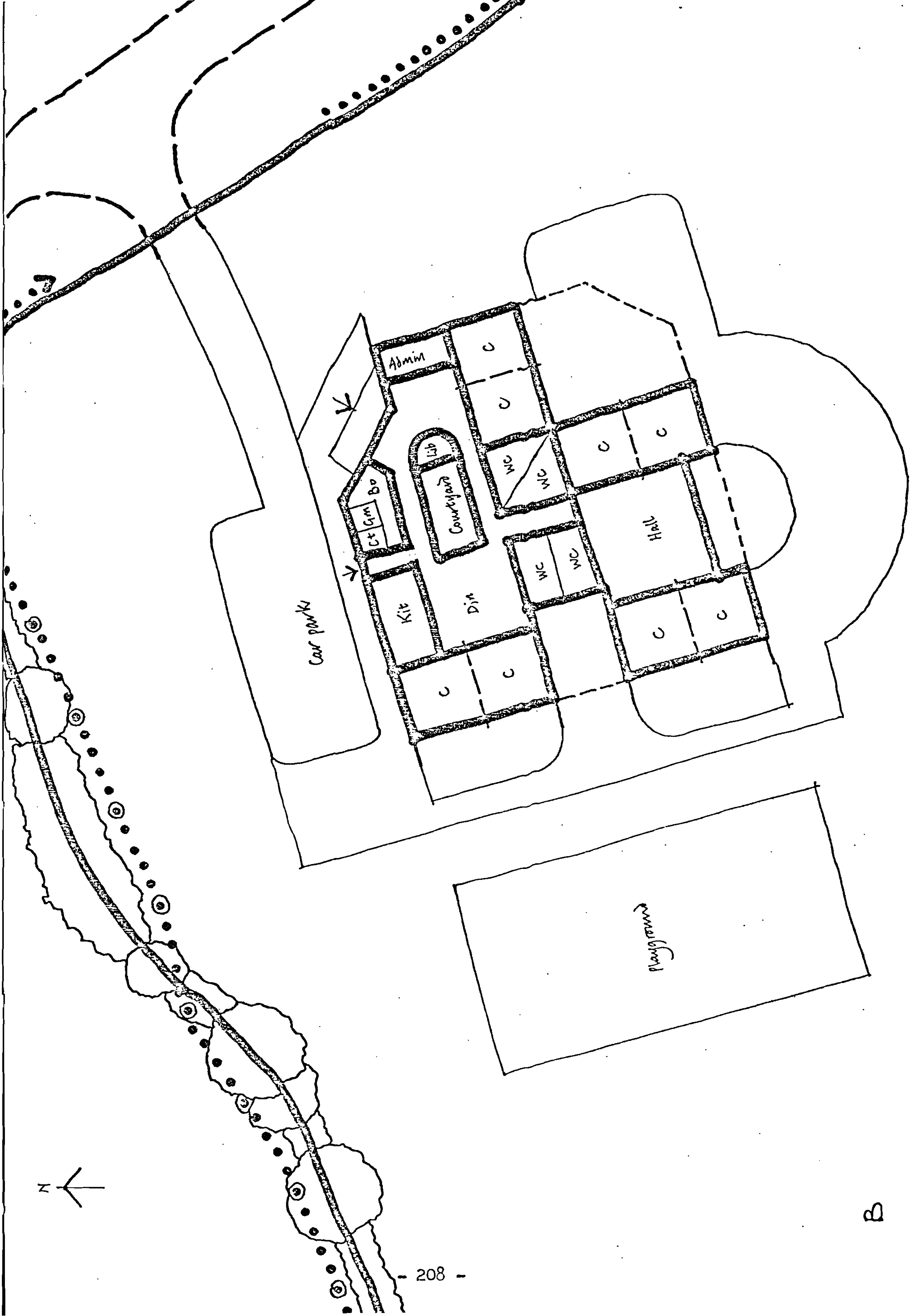


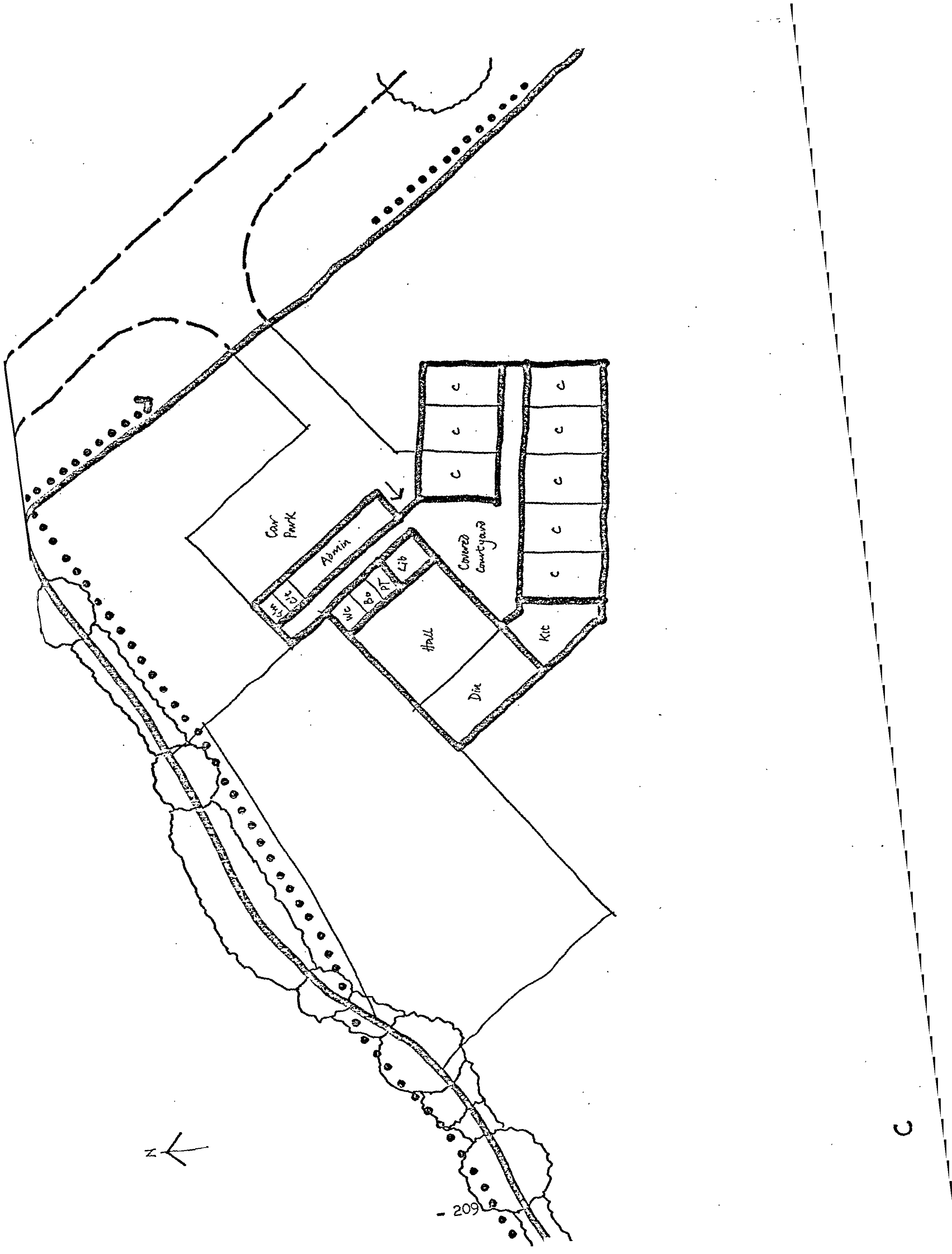
Key +—— Overall evaluation
×- - - Additive utility model indices

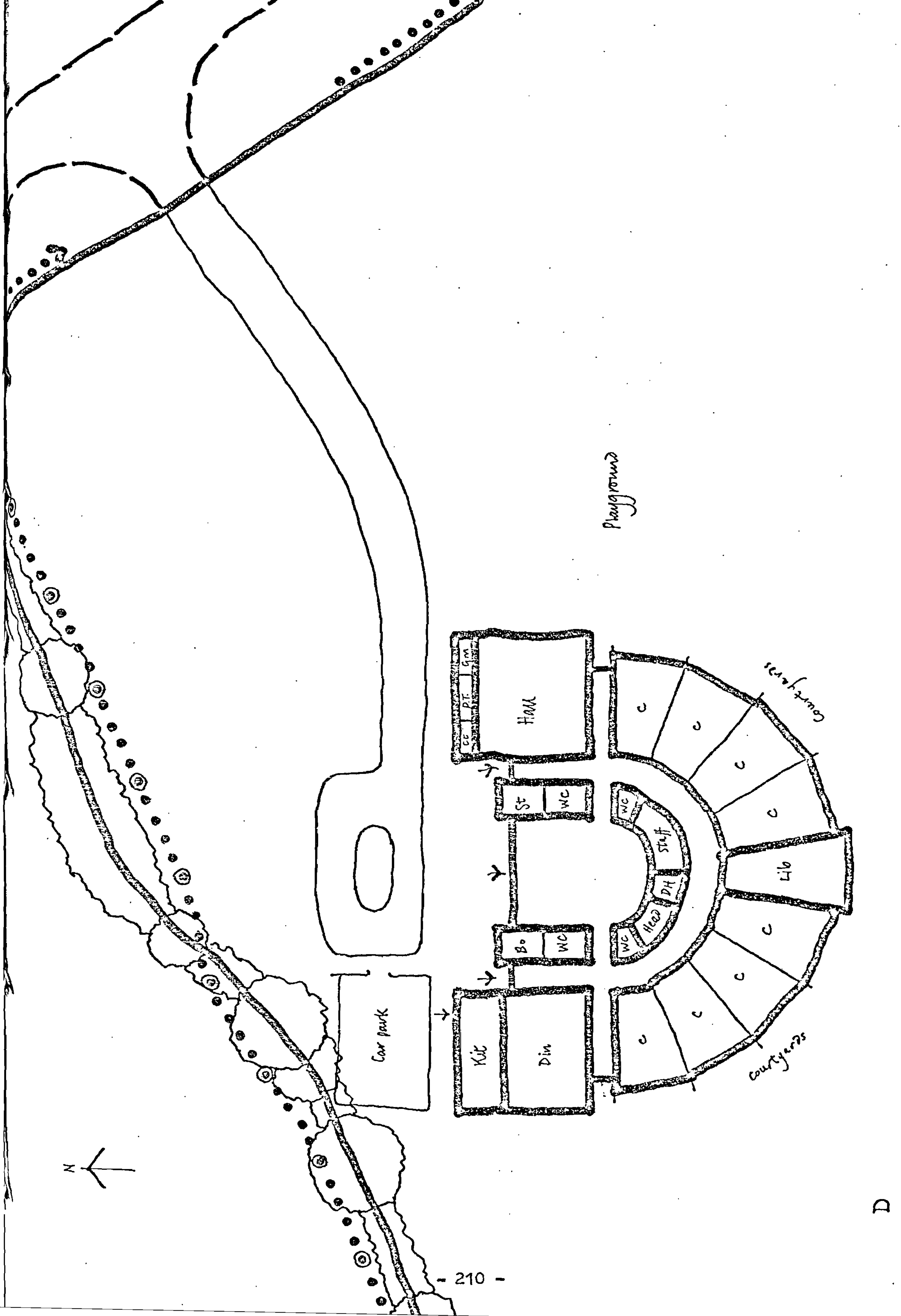
Figure 9.2 The redrawn school plans

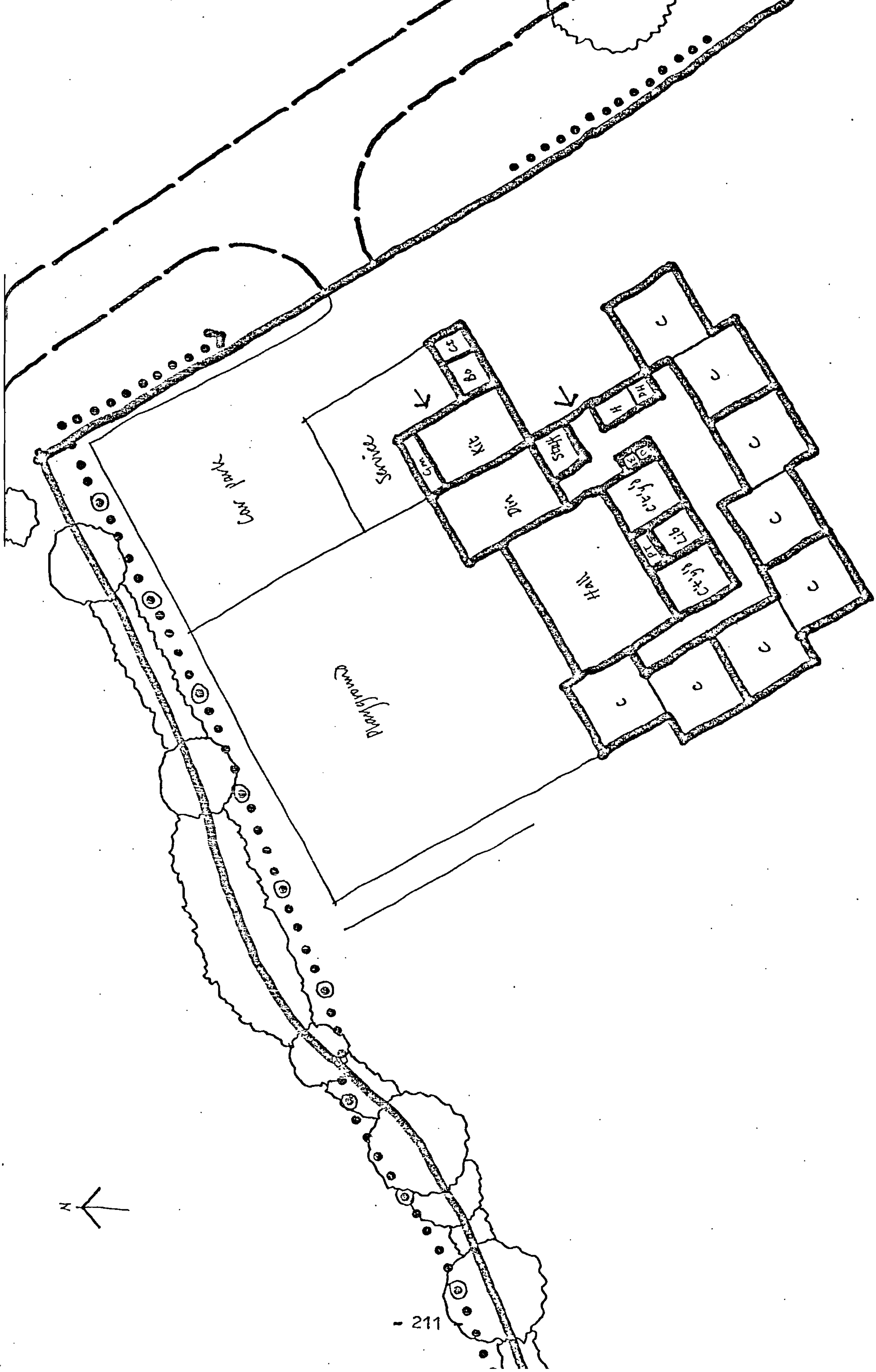
The following six pages show the redrawn school plans prepared in Experiment Five.

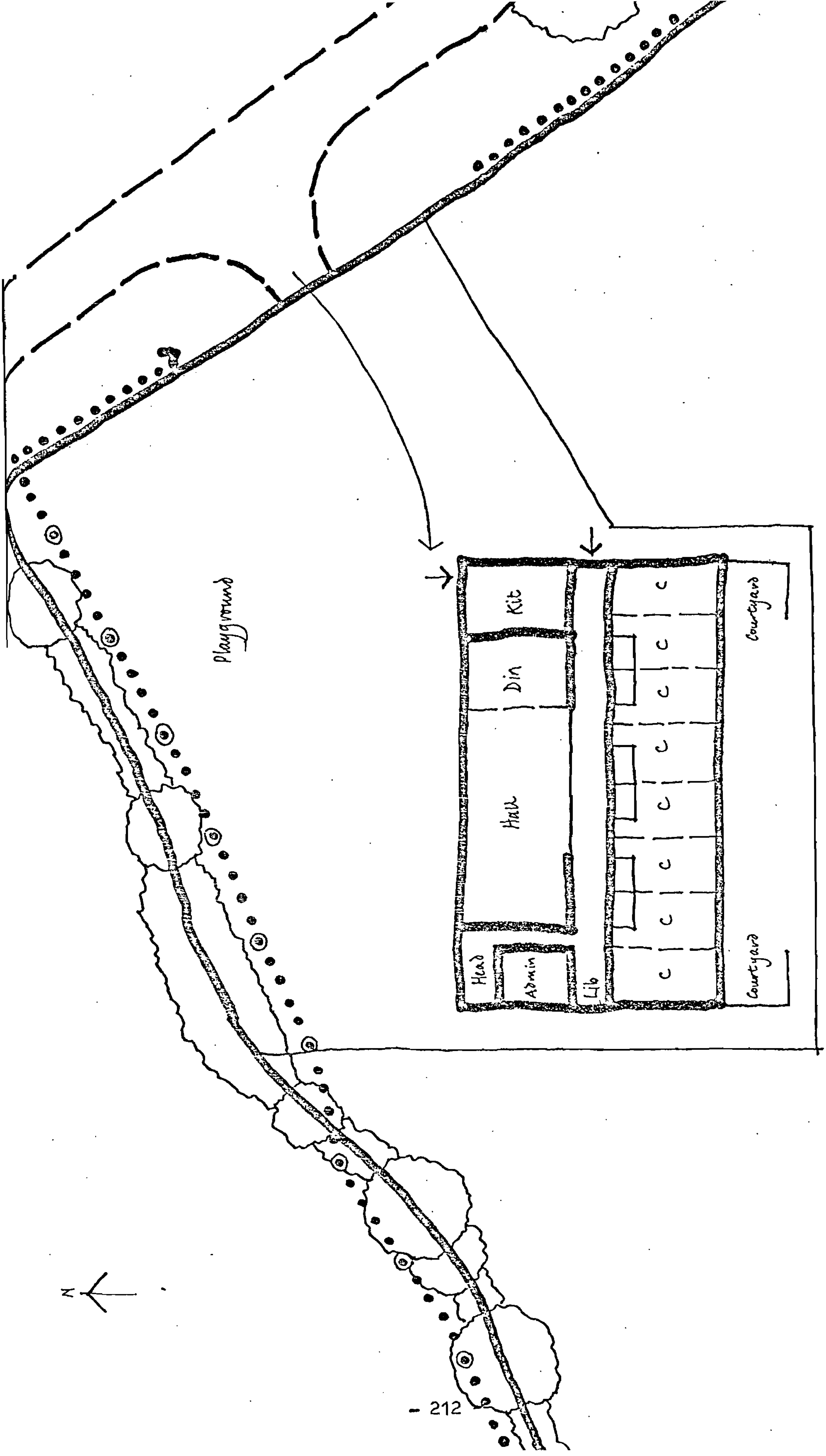

Admin. areas upstairs
wc
wc
Admin
Path
Car park
Servicing
Din
Hall (sunken)
Gallery
Lib
Courty'd
Courty'd
Kit
PT
Bo
Ct
Gm
c
A


Admin
C
C
WC
WC
Lib
Courtyard
Bo
Gm
Ct
Kit
Din
WC
WC
C
C
Hall
C
C
C
C
Car park
Playground
N
B


Car Park
Admin
Covered Courtyard
Lib
PT
Bo
WC
Hall
Din
Kit
C
C
C
C
C
C
C
C
N


Playground
Hall
CC
P.T.
Gm
St
WC
Bo
WC
WC
Staff
DH
Head
WC
Lib
C
courtyards
courtyards
Car park
Kit
Din
N
D


Car Park
Service
Kit
Din
Staff
Hall
Lib
Playground


Playground
Kit
Din
Hall
Head
Admin
Lib
C
Courtyard
Courtyard
N

CHAPTER 10 SUMMARY OF FINDINGS AND DIRECTIONS FOR FUTURE WORK

10.1 Validation of the Theory

The tentative theory proposed in chapter 2 was derived primarily from the literature search, and in drawing conclusions on the validity of the theory certain assumptions derived from the literature remain assumptions; they have not been tested in this thesis. This will become clear as conclusions are summarised, with reference to reiteration of the theory proposed in 2.7.

1 Designers may initially rely on a small set of strongly valued attributes to generate their design conjectures.

The designers had no difficulty in the experiments in offering a short list of valued attributes and proceeding to scale them.

In Experiments One and Five each subject when asked was able to describe six attributes for school planning with no difficulty. In the triadic comparisons in Experiments One and Five the designers did use between four and ten attributes to differentiate between the school plans, out of a possible twenty.

In Experiments Three and Four the designers in offering attributes by brainstorming gave initial lists of 35 and 19 respectively but found it quite feasible to reduce this number down to a basic list of six. In evaluation they were given the opportunity of describing up to twenty attributes in the triadic comparisons but four out of the twelve subjects used only the six attributes agreed at the start of the experiment and the rest used between four and nine.

However it has to be stated that the whole experiment, particularly the nature of the school plan as a design exercise, did encourage the use of only a limited number of attributes by the subjects. Similarly the nature of the way prioritization was used in the experiments also presupposed that a small number of attributes would account for designers' preferences. The supporting evidence for these assumptions comes only partially from the experiments, and primarily from the literature, as described in chapter 2.

2 <u>The attributes designers value may be understood as being weighted and ranked.</u>

The designers were able to ascribe weights to the attributes for design using the techniques employed in the experiments. They did not appear to have difficulty in doing so. In particular the consistency measure given by the maximum eigenvalue showed that, with few exceptions, they were able to perform this task with a considerable degree of internal consistency. This supports strongly the belief that designers do have an internally consistent value system which can be elicited verbally and numerically. There seem no reasons to believe that this value system does not guide and justify decision making in the design process. In the evaluation of alternative designs, the designers were again able to ascribe weights and to rank the attributes they were using.

3 <u>The weighting and ranking of attributes may change during the design process as the designer finds he needs to negotiate a solution. The internal consistency with which attributes are scaled may improve as a result of the design process.</u>

The results of Experiments Three, Four and Five are most pertinent to this question. Control parts of Experiments Four and Five, in which designers rated attributes for a holiday companion, showed that there were no arbitrary changes in attributes for a holiday companion during the experiment. Thus any changes in the rating of attributes for a school plan may be ascribed to the effect of the design process.

The results show that in the case of some subjects priorities did change, in others they did not. If it can be accepted that these three experiments, although slightly different in their respective organisation, yield results which are broadly comparable, then some quite interesting conclusions may be drawn.

In Experiment Three four out of six non-architects changed their priorities. In Experiment Five no architects changed their priorities significantly. These are very small samples to draw any conclusions from but it does seem possible that architects have a fairly fixed view of priorities; they know the major trade-offs they will make. Non-architects, given a school to design, may be barely able to predict what trade-offs they will make during the design process, so the design process may cause some of them to change their priorities. Students of architecture may be somewhere between these two states. In an unusual design problem using novel technology (Experiment Two) changes in priorities may be more pronounced, although there was a longer time interval between the two performances of prioritization, which might account for the recorded changes.

A feasible alternative explanation might appear to be that the architects used their own individual attributes in contrast with the non-architects and students who agreed sets of common attributes, and that this difference in experimental organisation accounts for the measured changes. This explanation is rejected however for two reasons. First in scaling attributes the students in Experiment Four achieved just as good a degree of internal consistency as the architects. Second the attributes for a holiday companion were agreed in just the same way as those for a school plan so if the changes in priorities were a result of the attributes having been agreed in a group as opposed to being elicited from the individual, the students would show a change in priorities among attributes for a holiday companion. This is not the case, and the alternative explanation is rejected.

No significant changes in the consistency with which subjects scaled attributes before and after design were recorded in the experiments.

4 <u>Designers may differ in the attributes they value, and in their evaluation of the same attributes.</u>

The first part of this statement may be considered from the findings of Experiment Five. The six architects were each asked for attributes they considered important in the planning of a school. While there are some common concerns, for example sunshine, light and orientation were mentioned by five out of six subjects, some attributes were mentioned by only one architect, for example simple structure/building system. There was even strong disagreement over some attributes: S5.4 mentioned 'Indication of hierarchy - head and teachers at the centre of the school - clarity of control' whereas S5.6 preferred to see 'Administrative hierarchy <u>not</u> reflected in plan arrangement'. Similar results are to be found in Experiment One. Designers do differ in the attributes they value.

For the purposes of Experiments Three and Four sets of common attributes were agreed in group situations, while in Experiment Two the attributes were decided by tutors. These experiments allow the second part of the statement to be considered. In Experiments Two and Four, in both of which the subjects were students of architecture, there was significant concordance between the ranks of attributes. These results are slightly surprising but may be explained by a theory put forward by March & Simon (1958) that the decision maker's organisational environment influences his value system. Two groups of subjects, each homogeneous in its composition of students of architecture, seem to have resulted in agreement about priorities.

In Experiment Three the subjects did not share similar backgrounds. Although before design there was agreement among priorities, changes caused by the design process resulted in significant differences among priorities after design. These designers differed in their evaluation of the same attributes.

From these results it appears that while designers may differ in the attributes they value, when a homogeneous group of designers agrees a set of common attributes there may be significant concordance among their ratings of attributes.

5 <u>The differences between designers' value systems will account in part for the differences between their design proposals.</u>

If Rescher's description of values were followed this statement would be a presupposition: if values are inferred from words and actions then words and actions would lead to assumptions about values in a decidedly non-empirical way. What is intended in the present experiments on the other hand is to try to show how verbally stated decisions about priorities are expressed as design proposals. Experiment Five provides the most dramatic demonstration. Each scheme shows quite clearly how the designer's value system manifested itself in the design. S5.1, for example, planned his school with the assembly hall in the centre, the library beside the main circulation route, the classrooms clustered around the courtyards, and the administration offices secluded by being on a separate first floor. This corresponds closely with what might be expected from his verbally expressed attributes. S5.6 planned quite a different school. A simple structural system, flexibly planned classes, and central circulation spine are clearly visible in the plan, and were among his verbally stated attributes.

One potential misunderstanding in this section should be cleared up. It concerns the results of Experiment Four. In this experiment there was statistically significant concordance between the six designers' ratings of attributes. However that does not imply that the six designers would therefore be expected to show no differences between their designs. The point is that although as a group there might be statistically significant concordance there will still be individual differences. Thus where, for example, S4.2 rated orientation as the prime attribute, S4.3 rated it only fifth. Differences of this kind, although not measured by the statistical procedures used, are nevertheless influential in their effects on the design process.

6 The differences between designer's value systems will account for their differing evaluations of alternative design proposals; they will favour plans which reflect their own priorities and reject plans which do not.

The two parts of this statement need to be separated. For it is an assumption of multi-attribute utility analysis that decision makers do make decisions to maximise their value or utility function, that is, that they do prefer plans which exhibit a high degree of fulfilment of the attributes which they value. This is therefore an assumption built in to the present experiments.

What is more important is to be able to examine and to explain the designer's different ratings of alternative designs, in terms of their differing value systems. This, the present experiments do. In evaluation the designers have generally been in agreement about the degree to which plans fulfil individual attributes, that is, they have agreed on the relative compactness, small scale, and so on, of the plans. Where they have differed about the relative importance of these attributes they have also differed in their evaluations of alternatives; where there has been agreement about the relative importance of the attributes there has been agreement about the rating of the designs.

10.2 Success of Prioritization as an Experimental Technique

Saaty's method for eliciting and scaling judgements has proved to be extremely valuable in the present thesis. The ability of the technique both to enable weights to be ascribed to activities, and to provide a measure of the internal consistency with which this has been done, allow designers' value systems as expressed verbally and numerically to be compared with their designs.

In evaluation, the same technique, together with the method of triadic comparisons and hierarchical cluster analysis, provided an efficient means of evaluating alternatives. Through comparisons of evaluations achieved by the overall merit ratings of designs, the hierarchical clustering, and the additive utility model indices, it has been possible to comment on the apparent reliability of the methods used. Figures 7.1, 8.1 and 9.1 show the comparisons in graphical form. The most encouraging sets of results are those in chapter 9 (Experiment Five) where as figure 9.1 shows correspondence between the three sets of results was generally good. The reason for this seems to be that the subjects were able to use their own attributes in the partial judgements, so the additive utility model indices better reflected their subjective judgements and therefore corresponded more to their overall evaluations. Where in Experiments Three and Four there were only low degrees of correspondence between the overall evaluation and the indices, it was possible for the most part to account for this within the experiments by showing that some subjects had not in fact used the six given attributes in their overall evaluations. This last finding shows how in these two experiments the techniques were used to permit integral checks to account for potential discrepancies in the results, over and above that provided by the measure of consistency in prioritization.

When subjects did use only their own attributes in the evaluation, Saaty's prioritization and the use of an additive utility model generally did provide consistent results to account for designers' evaluations in terms of their value systems. It could be argued that for such results to be achieved using the technique for the first time and without feedback of the answers or the opportunity to revise them is indicative of their potential. Nevertheless results of this kind prompt two arguments. The first is that if only half the subjects achieved meaningful results then is there any point in using techniques of this kind: would not an intuitive judgement be just as worthwhile? The counter argument is that if judgements demonstrate lack of correlation in their answers then should not systematic techniques be introduced to encourage rational thinking?

Research has shown that people make judgements which are inherently inconsistent, but that they are prepared to amend their judgements when the inconsistencies are revealed to them. This has been demonstrated, for example, about judgements based on statistical knowledge; even those with training in statistics made inaccurate judgements and were prepared to amend them when discrepancies were pointed out (Tversky & Kahneman, 1977). These findings are strongly supportive of the latter argument given above, in favour of the use of weighting techniques.

If it is accepted that weighting techniques are useful, then lack of internal consistency in the results necessitates the decision maker being given this as feedback in order for him to see his inconsistencies and to revise his judgements. It is in this context that these measurement techniques would become design aids. A dialogue takes place when the respondent is presented with the analysis of his judgements and is shown discrepancies and inconsistencies. Through this feedback mechanism he learns about his own subjective judgements in a way, perhaps, that simple introspection cannot reveal. Although that step has not been taken in the present thesis, these experiments have nevertheless revealed the potential use of these techniques as design aids.

10.3 Potential Use of Prioritization as a Design Aid

The relative ease of using the techniques for eliciting and structuring subjective data, and their demonstrated success in investigating designers' values and the evaluation of designs, indicate that they might prove useful as design aids. The potential range of uses includes the following:

1 Individual designers

In view of the results achieved by Abercrombie (reported in 1.1) that judgement may be improved by increasing self-awareness, the individual designer might use the design aid either to measure his own value system or to evaluate formally his alternative sketch plans for a design problem.

2 Designer teams

Teamwork in building design has become increasingly necessary, as reported in 1.1. Each member of the team may value different attributes, resulting in conflict of opinion about what the objectives in a particular design problem are. The use of the present technique, by eliciting the various value systems, would encourage discussion to resolve conflicting priorities among team members.

3 Evaluation of alternative design proposals

Alternative design proposals are commonly evaluated in planning departments, by companies commissioning buildings, in architectural competitions, and in schools of architecture. In all cases a formal evaluation method of this type would enable decisions to be justified by a logical and coherent set of principles, preventing the accusation of arbitrariness (Daley, 1969).

4 Teaching design

Judgements have been shown to be an essential part of the design process. The techniques used here provide a logical framework for making judgements explicitly. One of the aims of teaching design must be to encourage self awareness of the design process by the designer. Another must be to encourage him to think about alternative proposals before making final decisions. One point of being a student of architecture is to be able to explore a much wider range of alternatives in a way often precluded by pressures in design offices. A design aid of this kind encouages self awareness and provides a means for exploring the judgements implied by alternative designs.

5 Computer aided building design

One of the trends in CABD is to provide quantitative evaluations of the performance of design proposals. Calculations of daylight factors, heat losses, energy consumption, and capital and running costs are available on CEDAR and ABACUS systems, for example. A major criticism frequently made in the correspondence columns of architectural journals is that quantitative aspects will become concentrated upon by designers using CABD systems at the expense of qualitative judgements. Even Weinzapfel (1973), writing about IMAGE, noted that it was possible for designers to become seduced by the machine and to fail to take account of its limitations. Maver however reports different findings.

> "Contrary to the fears of many architectural practitioners, the use of CAAD techniques focusses increased attention in subjective value judgements, rather than less. As the measurable attributes are made more explicit, the necessary value judgements are forced to the surface of the design activity and thereby, themselves become more explicit" (Maver, Smith, Watts & Aish, 1979).

The designer's task thus becomes one of balancing the measurable attributes among themselves, and with the qualitative attributes he values. The present technique could prove a means of integrating these judgements. In doing so it would provide a slightly different approach from that taken by Cakin, described in section 2.5. The performance profiles which Cakin presented to some of his experimental subjects took the form of histograms. Two important points about these histograms are, first, that they convey information only about the attributes which the experimenter choses to measure and to present, and second, that the form of presentation gives each attribute equal weighting. Thus the evaluator may not only tend to rely on just the attributes which have been presented, but also may be discouraged from weighting even these attributes differently.

A technique like the present one could provide an interesting addition to the use of performance profiles. Either the designer could be given the quantitative information and then proceed to use the weighting technique in the light of this information. Or by integrating the technique into a CABD system it might be possible for the designer to make judgements about the relative importance of attributes in such a way that the performance profile could both show the attributes with which the designer was most concerned, and present them in a way which reflected his subjective weighting of attributes. Alternatively of course it would be possible to ascribe weights to the attributes not according to the designers' values but according to those of the client or of the users.

6 Clients' and users' values

If designers' value judgements have such a significant and demonstrable effect on their preferences among alternatives, and on their design proposals (as implied in this thesis and described by Darke, 1979, and Campbell, 1972) then the question is prompted as to how well these values correspond with those of their clients and of users of buildings. A great deal of research on the evaluation of the built environment has been carried out under the heading of architectural psychology but there is little evidence to suggest that the findings have yet been applied to a significant degree in design practice. Canter (1977a) emphasises the difference between the research findings of scientists and the assimilation and application of their results in practice.

On the other hand Abel (1975) reports the use of repertory grid technique, in a school of architecture, used to encourage students to appreciate the constructs of their clients. The author has used the present technique successfully to elicit the priorities of client representative in commissioning the design of a

> building, as part of a research project into industrial buildings (Nixon, Perera & Goumain, 1979). It may be the case that a technique of this kind could be used by designers to elicit the judgements of their clients, both to find their priorities and to obtain their reactions to alternative sketch proposals. This latter is the present direction of third generation design methods (Broadbent, 1977); those affected by design proposals are encouraged to evaluate the expert designers' conjectures.

This brief survey is not intended to convey the opinion that multi-attribute utility analysis is a universal tool, but at the same time as judgement is believed to be an essential part of the design process (Collins, 2.5) and as better judgements are believed to result from self awareness (Abercrombie, 1.1), so techniques of judgement analysis are highly relevant to the design process.

10.4 Concluding Remarks

Design is a complex activity, and one which needs to be simplified in order to be understood and described. This thesis has based the simplification on value theory, taking as its major premise the view expressed most succinctly by March

> "making decisions with respect to matters of value *is* designing" (March, 1976b).

Existing observations of design, particularly those resulting from monitoring and interviewing designers, indicate the important role of value judgements in design. Value theory provides a framework for creating from these observations a tentative theory. The application of multi-attribute utility analysis, in the from of a scaling technique developed by Saaty, provides a means to elicit and to analyse designers' judgements. A series of experiments enable the hypotheses, which arise from the tentative theory, to be tested using the scaling technique.

The results of the experiments provide strong evidence in support of the tentative theory. They underline the virtue of value theory as a means to understanding design. They stress the importance of the subjective structuring of design problems. The success with which the scaling technique can be used to elicit this subjective structuring shows how frequently covert and implicit value judgements may be investigated experimentally, for the benefit of the design community, and practically for the self awareness of individual designers.

The implications of the results for teaching are that alternative value systems be explicitly discussed and compared, and that awareness be developed of the design implications of alternative value systems. For practice the findings imply that for designers to respond to clients' needs it is essential for them to be briefed as precisely as possible, or to find out as much as possible, about their clients' values. Alternatively if it should prove to be the case that designers' values remain unmodified by their clients' needs, then clients must choose designers whose values correspond to their own. Finally for research there seem to be several potential lines of development of these ideas: using judgement analysis techniques in conjunction with quantitative evaluations provided by CABD systems, exploring the effect of specifying objectives or values on design proposals, and exploring whether and how a designer's value system is expressed in several different design problems. In experiments to monitor design processes it seems essential to study strategies in the context of value systems and design proposals; if the differences between value systems have such obvious effects on designs, it seems highly likely that they will also affect designers' strategies. One or more of these studies would seem to be an important next step for design research.

APPENDICES

Appendix 4.1 The numerical scale for prioritization

THE SCALE AND ITS DESCRIPTION

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Weak importance of one over another	Experience and judgement slightly favour one activity over another
5	Essential or strong importance	Experience and judgement strongly favour one activity over another
7	Demonstrated importance	An activity is strongly favoured and its dominance is demonstrated in practice
9	Absolute importance	The evidence favouring one activity over another is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values between the two adjacent judgements	When compromise is needed

Note: This scale was given to subjects as Table 2 (see appendix 4.2).

Appendix 4.2 The form for scaling attributes

Name: T:

Before you are six attributes on lettered cards, in alphabetical order.

You are to take the cards two at a time in the order shown in table 1. (Do not look at the remaining four while considering each pair.) For each pair look at the scale in table 2. Decide which one, if either, of the two is the more important in the planning of a two-form entry primary school (to satisfy the given brief, attached) and underline it in table 1. In the column labelled 'score' allocate a number to that more important attribute, in accordance with the scale given. If you consider both to be equal do not underline either, and allocated a score of '1' to that pair (as described in the scale in table 2). Finally in the section 'overall rank' please give a rank ordering of the attributes in descending order of importance (i.e. most important first).

Please take as much time as you need. When you finish, bring this page to me.

Table 1

Pair (underline the more important attribute)		Score (using scale in table 2)
a	b	
b	c	
c	d	
d	e	
e	f	
a	c	
b	d	
c	e	
d	f	
a	d	
b	e	
c	f	
a	e	
b	f	
a	f	
Overall rank:		

Appendix 4.3 The form for scaling plans with respect to one attribute

Name: T:
Attribute:

Before you are six school plans on lettered cards, in alphabetical order.

You are to take the plans two at a time in the order shown in table 1. (Do not look at the remaining four while considering each pair.) For each pair look at the scale in table 2. Decide which one of the school plans would be the better with respect only to the attribute above, and underline its letter in table 1, column 1. In the column labelled 'score' (table 1, column 2) allocate a number to the better school plan relative to the other of the pair, in accordance with the scale given in table 2. If you consider both plans to be equally good do not underline either, and allocate a score of '1' to that pair (as described in the scale in table 2). Finally in the section 'overall rank' give a rank ordering of the plans with respect to the attribute above in descending order (i.e. best first).

Please take as much time as you need. When you finish, bring this page to me.

Table 1

Pair (underline the better plan)		Score (using scale in table 2)
A	B	
B	C	
C	D	
D	E	
E	F	
A	C	
B	D	
C	E	
D	F	
A	D	
B	E	
C	F	
A	E	
B	F	
A	F	
Overall rank:		

Appendix 4.4 The form for scaling plans for their overall merit

Name: T:

Before you are six school plans on lettered cards, in alphabetical order.

You are to take the plans two at a time in the order shown in table 1. (Do not look at the remaining four while considering each pair.) For each pair look at the scale in table 2. Decide which one of the school plans would be the better school and underline its letter in table 1, column 1. In the column labelled 'score' (table 1, column 2) allocate a number to the better school plan relative to the other of the pair, in accordance with the scale given in table 2. If you consider both plans to be equally good do not underline either, and allocate a score of '1' to that pair (as described in the scale in table 2). Finally in the section 'overall rank' please give a rank ordering of the plans in descending order (i.e. best first)

Please take as much time as you need. When you finish, bring this page to me.

Table 1

Pair (underline the better plan)		Score (using scale in table 2)
A	B	
B	C	
C	D	
D	E	
E	F	
A	C	
B	D	
C	E	
D	F	
A	D	
B	E	
C	F	
A	E	
B	F	
A	F	
Overall rank:		

Appendix 4.6 The form for triadic comparisons (1) where attributes given

Name: T:

Before you are six school plans on capital-lettered cards, and a list of attributes identified by lower-case letters.

Consider the plans three at a time in the order shown below. For each triad separate out a pair which have or do not have an important attribute in common, which makes them similar and which differentiates them from the third plan. The attribute may be, but need not be, one of the six given.

When you have decided:

1 Underline the pair you judge similar in column 1.

2 Write the attribute in column 2.

3 In column 3 write down the letters of the plans which do demonstrate this attribute; it may be the pair which both demonstrate it, or the different one, and you must specify which.

It is recommended that for each set of three you look at alternative ways of pairing off two before making a final decision.

Please take as much time as you need. When you finish bring this page to me.

Triad (underline similar pair)			Attribute	Plans fulfilling attribute
A	B	C		
A	D	E		
B	D	F		
B	C	E		
A	E	F		
A	C	D		
B	E	F		
A	C	F		
C	D	E		
A	B	D		
C	E	F		
B	C	D		
A	D	F		
A	B	E		
C	D	F		
A	B	F		
B	D	E		
A	C	E		
B	C	F		
D	E	F		

Appendix 4.6 The form for triadic comparisons (2) to elicit attributes

Name: T:

Before you are six school plans on capital-lettered cards.

Consider the plans three at a time in the order shown below. For each triad separate out a pair which have or do not have an important attribute in common, which makes them similar and which differentiates them from the third plan.

When you have decided:

1 Underline the pair you judge similar in column 1.

2 Write the description of the attribute, briefly, in column 2.

3 In column 3 write down the letters of the plans which do demonstrate this attribute; it may be the pair which both demonstrate it, or the different one, and you must specify which.

It is recommended that for each set of three you look at alternative ways of pairing off two before making a final decision.

Please take as much time as you need. When you finish, bring this page to me.

Triad (underline similar pair)			Attribute	Plans fulfilling attribute
A	B	C		
A	D	E		
B	D	F		
B	C	E		
A	E	F		
A	C	D		
B	E	F		
A	C	F		
C	D	E		
A	B	D		
C	E	F		
B	C	D		
A	D	F		
A	B	E		
C	D	F		
A	B	F		
B	D	E		
A	C	E		
B	C	F		
D	E	F		

Appendix 4.7 Computer program to calculate additive utility model indices and hierarchical cluster analysis

```
10 REM ADDITIVE UTILITY MODEL
20 PRINT "INPUT NO OF STIMULUS ITEMS"
21 LPRINT "INPUT NO OF STIMULUS ITEMS"
30 INPUT N
31 LPRINT N
35 PRINT:PRINT:PRINT
36 LPRINT:LPRINT:LPRINT
40 DIM A(10,10)
50 FOR I=1 TO N
60 FOR J=1TON
70 A(I,J)=0
80 NEXT J
90 NEXT I
100 FOR X=1 TO N-2
110 FOR Y=X+1 TO N-1
120 FOR Z=Y+1 TO N
130 PRINT "COMPARE ITEMS" X,Y,Z
131 LPRINT"COMPARE ITEMS",X,Y,Z
140 INPUT I,J
146 LPRINT I,J
150 A(I,J)=A(I,J)+1
160 A(J,I)=A(I,J)
170 NEXT Z
180 NEXT Y
190 NEXT X
193 PRINT:PRINT:PRINT
194 LPRINT:LPRINT:LPRINT
200 FOR I=1 TO N
210 FOR J=1 TO N
220 PRINT A(I,J),
225 LPRINT A(I,J),
230 NEXT J
231 PRINT
232 LPRINT
235 NEXTI
240 GOTO1010
241 PRINT:PRINT:PRINT
242 LPRINT:LPRINT:LPRINT
245 PRINT"SIMILARITY"
246 LPRINT"SIMILARITY"
247 PRINT"LEVEL"
248 LPRINT"LEVEL"
250 FOR P=N-2 TO 1 STEP -1
260 FOR I=1 TO N
270 FOR J=I+1 TO N
280 IF A(I,J)<>P GOTO 320
310 PRINTP;:PRINTTAB(I*10):PRINTI;:PRINTTAB(J*10):PRINTJ
315 LPRINTP;:LPRINTTAB(I*10):LPRINTI;:LPRINTTAB(J*10):LPRINTJ
320 NEXT J
330 NEXT I
335 PRINT:PRINT:PRINT
336 LPRINT:LPRINT:LPRINT
340 NEXT P
350 GOTO4080
```

```
1010 REM
1020 PRINT: PRINT: PRINT
1030 N1=N
1040 DIM B(10,10),X(10),Y(10)
1050 PRINT "INPUT SCALED JUDGEMENTS OF ITEMS"
1051 LPRINT"INPUT SCALED JUDGEMENTS OF ITEMS"
1060 FOR K=N-1TO 1 STEP -1
1062 FOR I=1 TO K
1064 J=I+N-K
1066 PRINT"COMPARE ITEMS",I,"AND",J
1067 LPRINT "COMPARE ITEMS";I;"AND";J
1068 INPUT B(J,I)
1069 B(I,J)=1/B(J,I)
1070 NEXT I
1072 NEXT K
1074 FOR I=1 TO N
1076 B(I,I)=1
1078 NEXT I
1110 FOR J= 1 TO N
1120 FOR I=1 TO N
1125 PRINTTAB(I*8-8)
1126 LPRINTTAB(I*8-8)
1130 PRINTB(I,J);
1131 LPRINTB(I,J);
1150 NEXT I
1155 PRINT
1156 LPRINT
1160 NEXT J
1165 PRINT: PRINT
1166 LPRINT:LPRINT
1200 S=0
1210 FOR I=1TON
1220 FOR J=1TON
1230 S=S+B(I,J)*B(I,J)
1240 NEXT J
1250 NEXT I
1270 S=SQR(S)
1280 SF=S*1E-06/N
1290 FOR J=1TON
1300 X(J)=0
1310 NEXT J
1320 X(N)=1
1330 YB=10
1340 FOR I=1TON
1350 Y(I)=0
1360 FOR J=1TON
1370 Y(I)=Y(I)+B(J,I)*X(J)
1380 NEXT J
1390 NEXT I
1400 FOR I=1 TO N
1410 X(I)=Y(I)/Y(N)
1420 NEXT I
1430 D=ABS(Y(N)-YB)
1440 IFD-SF<=0THEN 1465
1450 YB=Y(N)
```

```
1460 GCTO 1340
1465 PRINT"OVERALL WEIGHTING OF ITEMS"
1466 LPRINT"OVERALL WEIGHTING OF ITEMS"
1470 PRINT"MAX.  EIGENVALUE:",Y(N)
1471 LPRINT"MAX.  EIGENVALUE:",Y(N)
1480 NO=0
1490 FOR J=1 TO N
1500 NO=NO +X(J)
1510 NEXT J
1520 PRINT 'CORRESPONDING EIGENVECTOR      NORMALISED EIGENVECTOR"
1530 FOR I=1 TO N
1540 PRINT X(I),"                          "X(I)/NO
1541 LPRINTX(I),"                     ",X(I)/NO
1545 P(I)=X(I)/NO
1550 NEXT I
1560 PRINT:PRINT:PRINT
1570 LPRINT:LPRINT:LPRINT
2010 REM TO WEIGHT ATTRIBUTES
2030 PRINT "INPUT NO OF ATTRIBUTES"
2031 LPRINT"INPUT NO OF ATTRIBUTES"
2032 INPUT N2
2033 LPRINT N2
2040 DIMC(10,10),V(10),W(10)
2050 PRINT "INPUT SCALED JUDGEMENTS OF ATTRIBUTES"
2060 FOR K=N2 -1  TO1 STEP -1
2062 FOR I=1 TO K
2064 J=I+N2-K
2066 PRINT"COMPARE ATTRIBUTES";I;"AND";J
2067 LPRINT"COMPARE ATTRIBUTES";I;"AND",J
2068 INPUT C(J,I)
2069 C(I,J)=1/C(J,I)
2070 NEXT I
2072 NEXT K
2074 FOR I=1 TO N2
2076 C(I,I)=1
2078 NEXT I
2110 FOR J=1 TO N2
2120 FOR I=1 TO N2
2125 PRINTTAB(I*8-8)
2126 LPRINTTAB(I*8-8)
2130 PRINT C(I,J);
2135 LPRINTC(I,J);
2150 NEXT I
2155 PRINT
2156 LPRINT
2160 NEXT J
2200 S=0
2210 FOR I=1 TO N2
2220 FOR J=1 TO N2
2230 S=S+C(I,J)*C(I,J)
2240 NEXT J
2250 NEXT I
2270 S=SQR(S)
2280 SF=S*1E-06/N2
2290 FOR J=1 TO N2
2300 V(J)=0
2310 NEXT J
2320 V(N2)=1
```

```
2330 WN=10
2340 FOR I=1 TO N2
2350 W(I)=0
2360 FOR J=1 TO N2
2370 W(I)=W(I)+C(J,I)*V(J)
2380 NEXT J
2390 NEXT I
2400 FOR I=1 TO N2
2410 V(I)=W(I)/W(N2)
2420 NEXT I
2430 D=ABS(W(N2)-WN)
2440 IFD-SP<=0THEN2465
2450 WN=W(N2)
2460 GOTO 2340
2465 PRINT"WEIGHTING OF ATTRIBUTES"
2466 LPRINT"WEIGHTING OF ATTRIBUTES"
2470 PRINT"MAX EIGENVALUE:",W(N2)
2471 LPRINT"MAX. EIGENVALUE",W(N2)
2480 NO=0
2490 FOR J=1 TO N2
2500 NO=NO+V(J)
2510 NEXT J
2520 PRINT "CORRESPONDING EIGENVECTOR       NORMALISED EIGENVECTOR"
2521 LPRINT"CORRESPONDING EIGENVECTOR       NORMALISED EIGENVECTOR"
2530 FOR I=1 TO N2
2540 PRINT V(I),"                              ",V(I)/NO
2541 LPRINTV(I),"                          ",V(I)/NO
2545 Q(I)=V(I)/NO
2550 NEXT I
2560 PRINT:PRINT:PRINT
2570 LPRINT:LPRINT:LPRINT
3040 FOR L=1 TO N2
3045 PRINT:PRINT:PRINT
3046 LPRINT:LPRINT:LPRINT
3050 PRINT "INPUT SCALED JUDGEMENTS OF ITEMS WITH RESPECT TO ATTRIBU
3051 LPRINT"INPUT SCALED JUDGEMENTS OF ITEMS WITH RESPECT TO ATTRIBU
3060 FOR K=N1-1 TO 1 STEP-1                                        [NO";L
3062 FOR I=1 TO K
3064 J=I+N1-K
3066 PRINT "COMPARE ITEMS",I,"AND",J
3067 LPRINT"COMPARE ITEMS";I;"AND";J
3068 INPUT D(J,I)
3069 D(I,J)=1/D(J,I)
3070 NEXT I
3072 NEXT K
3074 FOR I=1 TO N1
3076 D(I,I)=1
3078 NEXT I
3110 FOR J=1 TO N1
3120 FOR I=1 TO N1
3125 PRINTTAB(I*8-8)
3126 LPRINTTAB(I*8-8)
3130 PRINT D(I,J);
3131 LPRINT D(I,J);
3150 NEXT I
3155 PRINT
3156 LPRINT
3160 NEXT L
```

```
3200 S=0
3210 FOR I=1 TO N1
3220 FOR J=1 TO N1
3230 S=S+D(I,J)*D(I,J)
3240 NEXT J
3250 NEXT I
3270 S=SQR(S)
3280 SF=S*1E-05/N1
3290 FOR J=1 TO N1
3300 T(L,J)=0
3310 NEXT J
3320 T(L,N1)=1
3330 UN=10
3340 FOR I=1 TO N1
3350 U(I)=0
3360 FOR J=1 TO N1
3370 U(I)=U(I)+D(J,I)*T(L,J)
3380 NEXT J
3390 NEXT I
3400 FORI=1 TO N1
3410 T(L,I)=U(I)/U(N1)
3420 NEXT I
3430 D=ABS(U(N1)-UN)
3440 IFD-SF<=0THEN3465
3450 UN=U(N1)
3460 GOTO 3340
3465 PRINT"WEIGHTING OF ITEMS WITH RESPECT TO ATTRIBUTE NO";L
3466 LPRINT"WEIGHTING OF ITEMS WITH RESPECT TO ATTRIBUTE NO";L
3470 PRINT "MAX EIGENVALUE:",U(N1)
3471 LPRINT"MAX. EIGENVALUE:",U(N1)
3480 NO=0
3490 FOR J=1 TO N1
3500 NO=NO+T(L,J)
3510 NEXT J
3520 PRINT "CORRESPONDING EIGENVECTOR     NORMALISED EIGENVECTOR"
3521 LPRINT"CORRESPONDING EIGENVECTOR     NORMALISED EIGENVECTOR"
3530 FOR I=1 TO N1
3540 PRINT T(L,I),"                    ",T(L,I)/NO
3541 LPRINTT(L,I),"                    ",T(L,I)/NO
3545 R(L,I)=T(L,I)/NO
3550 NEXT I
3560 NEXT L
4000 FOR I=1 TO N1
4010 FOR L=1 TO N2
4020 S(L,I)=Q(L)*R(L,I)
4040 E(I)=E(I)+S(L,I)
4050 NEXT L
4060 NEXT I
4070 GOTO241
4080 PRINT"INDEX";
4081 LPRINT"INDEX";
4090 FORI=1TON
4100 PRINTTAB(I*10);
4101 LPRINTTAB(I*10);
4110 PRINT E(I);
4111 LPRINT E(I);
4120 NEXTI
4130 END
```

Appendix 4.8 Computer program to calculate Kendall's coeffieient of concordance

```
5 REM KENDALL
10 REM
20 PRINT "INPUT NO. OF JUDGES"
22 LPRINT "NO. OF JUDGES'
30 INPUT K :PRINT :LPRINT
35 PRINT
36 LPRINTK
40 PRINT "INPUT NO. OF ENTITIES"
50 INPUT N :PRINT :LPRINT
52 LPRINT "NO. OF ENTITIES"
55 PRINT
56 LPRINT N
60 DIM A(K,N),B(K,N), E(N),F(N),U(N),X(N),Y(N),Z(N)
70 FOR I=1TOK
75 L=0
80 PRINT "INPUT RANKS FOR JUDGE NO. ",I
82 LPRINT "RANKS FOR JUDGE NO. ",I
90 FOR J=1TON
100 INPUT A(I,J) :PRINT
102 LPRINT A(I,J)
103 L=L+A(I,J)
110 NEXT J
115 IF L<>(N+1)*N/2 THEN 116 ELSE 120
116 PRINT "ERROR IN RANKS FOR JUDGE NO. ",I
117 END
120 NEXT I
130 FOR J=1TO N
140 X(J)=0
150 NEXT J
160 FOR J=1TO N
170 FOR I=1TO K
180 X(J)=X(J)+A(I,J)
190 NEXT I
200 NEXT J
210 FOR J=1TO N
220 E(J)=K*(N+1)/2
230 NEXT J
270 FOR J=1TO N
280 Y(J)=X(J)-E(J)
290 NEXT J
300 FOR J=1TO N
310 Z(J)=Y(J)[illegible]
316 NEXT J
320 [illegible]
330 FOR J=1TO N
340 [illegible]
342 NEXT J
344 PRINT"SSUM OF SQUARES =",S
345 LPRINT "SUM OF SQUARES =",S
360 PRINT "ARE THERE MANY TIED RANKS? TYPE 1 FOR YES, 0 FOR NO"
370 INPUT C
```

```
380 IF C=1 THEN 1000 ELSE 390
390 W=12*S/((K*K)* (N*N*N-N))
400 PRINT "COEFFICIENT OF CONCORDANCE IS ",W
410 LPRINT "COEFFICIENT OF CONCORDANCE IS",W
420 GOTO1220
1000 FOR I=1TO K
1010 FOR J=1TO N
1020 B(I,J)=A(I,J)
1030 NEXT J
1040 NEXT I
1050 V=0
1060 FOR I=1TO K
1070 FOR Q=1TO N STEP .5
1075 T=0
1080 FOR J=1TO N
1090 IF B(I,J)<>Q THEN 1130
1100 T=T+1
1130 NEXT J
1135 V=V+(T*T*T-T)
1140 NEXT Q
1180 NEXT I
1190 W=12*S/((K*K)*(N*N*N-N)-K*V)
1200 PRINT "COEFFICIENT OF CONCORDANCE "
1205 LPRINT "COEFFICIENT OF CONCORDANCE"
1210 PRINT "CORRECTED FOR TIES IS" ,W
1215 LPRINT "CORRECTED FOR TIES IS",W
1220 D=N-1
1230 K2=K*W*(N-1)
1240 IF N>7 GOTO 1270
1250 PRINT"SEE SIEGEL P236 FOR SIGNIFICANCE OF W"
1255 LPRINT "SEE SIEGEL P236 FOR SIGNIFICANCE OF W"
1260 END
1270 PRINT"CHI SQUARE =",K2,"FOR",D,"DEGREES OF FREEDOM"
1275 LPRINT"CHI SQUARE =",K2,"FOR",D,"DEGREES OF FREEDOM"
1280 END
```

Appendix 4.9 Computer program to calculate maximum eigenvalue and normalised eigenvectors

```
5 REM SAATY
10 REM MAX EIGENVALUE BY ITERATIVE METHOD
20 PRINT:PRINT:PRINT
30 PRINT "INPUT NO OF STIMULUS ITEMS"
31 INPUT N
40 DIM A(10,10),X(10),Y(10)
50 PRINT "INPUT SCALED JUDGEMENTS"
51 PRINT
60 FOR K=N-1TO 1 STEP -1
62 FOR I=1TOK
64 J=I+N-K
66 PRINT "COMPARE ITENS ",I,"AND",J
68 INPUT A(J,I)
69 A(I,J)=1/A(J,I)
70 NEXT I
72 NEXT K
74 FOR I=1TON
76 A(I,I)=1
78 NEXT I
110 FOR J=1TON
120 FOR I=1 TO N
130 PRINT A(I,J).
150 NEXT I
155 PRINT
160 NEXT J
165 PRINT:PRINT
200 S=0
210 FOR I=1TON
220 FOR J=1TON
230 S=S+A(I,J)*A(I,J)
240 NEXT J
250 NEXT I
270 S=SQR(S)
280 SF=S*1E-06/N
290 FOR J=1TON
300 X(J)=0
310 NEXT J
320 X(N)=1
330 YN=10
340 FOR I=1TON
350 Y(I)=0
360 FOR J=1TON
370 Y(I)=Y(I)+A(J,I)*X(J)
380 NEXT J
390 NEXT I
400 FOR I=1TON
410 X(I)=Y(I)/Y(N)
```

```
420 NEXT I
430 D=ABS(Y(N)-YN)
440 IFD-SF<=0THEN470
450 YN=Y(N)
460 GOTO 340
470 PRINT "MAX EIGENVALUE :",Y(N)
471 PRINT:PRINT
480 NO=0
490 FOR J=1TON
500 NO=NO+X(J)
510 NEXT J
520 PRINT "CORRESPONDING EIGENVECTOR      NORMALISED EIGENVECTOR"
530 FOR I=1TON
540 PRINT X(I),"                              ",X(I)/NO
550 NEXT I
560 END
```

Appendix 5.1 The brief (Experiment One)

PROJECT: Proposed two-form entry County Primary School,
Nascot Wood Road, Watford, Hertfordshire.

Basic activities to be accommodated in following spaces:

1	Assembly hall
2	Teaching area
3	Dining area
4	Kitchen
5	Administration
6	Space for library, either centralised or dispersed
7	Cloakrooms, storage, lavs and circulation all implied in the above

Other information:

All services available on site

Pedestrian and vehicular circulation required on site

Car park

Deliveries

External plan pitches required

Future extensions not anticipated

Appendix 6.1 The brief (Experiment Two)

DESIGN OF COAST GUARD STATION

You have been commissioned by the Customs and Excise Department to prepare a design and detailed drawings for a two man Coast Guard Station to be erected on a number of sites in all parts of the UK.

The large number of units required (150 off) allows the economic use of grp and you are to design using this material. The enclosed lookout platform should allow 180° view and the viewer's eye level should be five metres above average ground level allowing for location on sloping sites.

In addition to the lookout platform you should provide a small office/telex room, galley and toilet. It will probably be necessary to provide a draught lobby. Parking area will be required near to the station for the coastguards' cars.

Detailed requirements for the lookout room are as follows:

- i bench space for charts, ordinance survey maps, plan chests and log books
- ii shelving space for books (say 3m)
- iii binoculars - preferably mounted on track from the ceiling
- iv telephone
- v notice boards
- vi VHF set
- vii Aldis lamps
- viii compass
- ix perspex covers to maps

You should also allow on the site for a signal mast (flags), base for firing mortars, standby generator, incinerator and septic tank.

Assume that reasonable road access is available to all sites. Other points for consideration are condensation, dimming of artificial lighting at night, rain on glazing, sun shading, mullions obstructing view, draught lobby, visibility from sea and demountability of unit.

Appendix 6.2 Timetable (Experiment Two)

Monday	10.30	Brief handed out. Introductory talk about current approaches in design research.
	12.00	Students departed. Tutors agreed four aspects of grp technology as foci for the exercise. Experimenter allocated students to teams and groups.
	14.00	Teams X and Y performed prioritization of attributes. Team Z absent.
	14.30	Students told allocation into groups A, B, C, D. Students began background investigation in groups.
Tuesday	10.00	Groups A and B presented findings.
	14.00	Groups C and D presented findings.
	15.30	Team X given results of prioritization.
Wednesday		-
Thursday	All day	Students design schemes individually.
	16.00	Students submit schemes.
Friday	10.30	All students (Teams X, Y and Z) perform prioritization.
	11.00	Schemes numbered randomly. All students mark all schemes. Experimenter begins analysing results.
	14.00	Students given feedback of some of their results as basis for discussion.

Appendix 6.3 The scale for marking design schemes (Experiment Two)

Overall range	1 to 20
1 to 4	Very poor
5 to 8	Below average
9 to 12	Average
13 to 16	Above average
17 to 20	Very good

Appendix 7.1 The brief (Experiments Three, Four and Five)

BRIEF FOR A TWO-FORM ENTRY PRIMARY SCHOOL

Schedule of Accommodation

1	Teaching areas: 8 classrooms @ $60m^2$, including associated wc's.
2	Assembly hall: $160m^2$
3	PT store, associated with assembly hall: $10m^2$
4	Dining area: $100m^2$
5	Kitchen: $60m^2$ with servery to dining room
6	Administration: Head teacher's room: $14m^2$ Deputy head's room: $11m^2$ Staffroom: $20m^2$ Stationery store: $10m^2$ Male cloakroom: $5m^2$ Female cloakroom: $5m^2$
7	Library: $20m^2$
8	Caretaker: $10m^2$
9	Groundsman: $10m^2$
10	One or two courtyards, total $75m^2$
11	Boiler room: $10m^2$

Total area of accommodation: $1000m^2$

Other information

1	All services available on site.
2	Pedestrian and vehicular circulation required on site.
3	Car park: $200m^2$
4	Hard playground: $1200m^2$, with direct access to assembly hall.
5	Future extension not anticipated.

Appendix 7.2 Time-table (Experiment Three)

Day One

09.45	Introduction
10.00	Discussion of attributes and prioritization
12.00	School plan design
13.00	Lunch
14.00	School plan design continues
16.00	Experimenter collects schemes and redraws
16.15	Prioritization of attributes

Day Two

09.30	Evaluation of designs

Appendix 7.3 List of attributes from brainstorming session (Experiment Three)

1	Away from busy roads
2	Road access to kitchens
3	Easy access to outside areas
4	Easy access to common areas
5	Views to countryside
6	Awareness of future developments
7	Heating system
8	Playground/vehicle separation
9	Single entrance/exit - for convenience
10	Close to housing
11	Small scale for small people
12	Drainage and site works required
13	Flexibly defined areas (small groups)
14	Wind direction
15	Sunlight/South light
16	Open space classrooms
17	Noise/environment (external)
18	Noise from playground to housing
19	Security from farm animals, dogs
20	Underground
21	Energy conservation
22	Budget
23	Soft edge hardware construction
24	Water supply in classrooms
25	Landscaping
26	No limit to courtyards
27	Encourage parental involvement
28	Swimming pool and plant
29	PT store/playing field link
30	Glare protection
31	Covered ways
32	Noise (internal environment)
33	Building regulations and other codes
34	Disabled

Appendix 8.1 Time-table (Experiment Four)

Day One

09.45	Introduction
10.00	Discussion of attributes and prioritization
12.00	School plan design
13.00	Lunch
14.00	School plan design continues
16.00	Experimenter collects schemes and redraws
16.15	Prioritization of attributes

Day Two

09.30	Evaluation of designs

Appendix 8.2 List of attributes for school planning from brainstorming session (Experiment Four)

1	Small scale (intimate)
2	Good thermal properties
3	Avoidance of long corridors
4	Acoustic properties: noise
5	Integration: non-hierarchical
6	Easily understood plan - navigable
7	Safety exits for fire
8	Flexibility
9	Vandal proof
10	Focal point
11	Look like a school
12	Anticipate parental involvement
13	Lighting
14	Outdoor area for each classroom
15	Element of danger
16	Friendly: non-institutional
17	Circulation
18	Evening/weekend use
19	Orientation

BIBLIOGRAPHY

ABEL, C. (1975) 'The Instructional Simulation of Client Construct Systems', paper to Architectural Psychology Conference, University of Sheffield.

ABERCROMBIE, M.L.J. (1960) The Anatomy of Judgement, London, Hutchinson.

ABERCROMBIE, M.L.J., HUNT, S. & STRINGER, P. (1969) Selection and Academic Performance of Students in a University School of Architecture, London, Society for Research into Higher Education.

ABERCROMBIE, M.L.J. & HUNT, S.M. (1977) '1960 - 1970: Ten Years of Development in a School of Architecture', Report, University College London.

ABERCROMBIE, M.L.J. & TERRY, P.M. (1977) 'A Contribution to the Psychology of Designing', Journal of Architectural Education, XXX, 4, 15 - 18.

ABERCROMBIE, S. (1978) 'How To Aim Accurately At The Minds Of Child Readers', Design, 354, 60 - 63.

ACKOFF, R.L. (1962) Scientific Method: optimising applied research decisions, New York, Wiley.

ACKOFF, R.L. (1974) Redesigning the Future, New York, Wiley.

ACKOFF, R.L. & EMERY, F.E. (1972) On Purposeful Systems, London, Tavistock Puliciations.

ADAMS, E.W. & FAGOT, R. (1967) 'A Model of Riskless Choice', in EDWARDS & TVERSKY, 1967.

AISH, R. (1974) An Analytical Approach to the Design of a Man-Machine Interface, PhD thesis, University of Essex.

AKIN, O. (1978a) 'How do Architects Design?' Report, Department of Architecture, Carnegie-Mellon University.

AKIN, O. (1978b) 'Exploration of the Design Process', Report, Department of Architecture, Carnegie-Mellon University (published in Design Methods and Theories, 13, 3/4 115 - 119 in 1979).

ALCOCK, D. (1977) Illustrating BASIC, Cambridge, Cambridge University Press.

ALEXANDER, C. (1964) Notes on the Synthesis of Form, Cambridge, Mass., Harvard University Press.

ALEXANDER, C. & MANHEIM, M. (1962) The Use of Diagrams in Highway Route Location: An experiment, Civil Engineering Systems Laboratory M.I.T. Publication no.161, Cambridge, Mass. (refered to in GRANT, 1974).

ALEXANDER, C. & POYNER, B. (1970) 'The Atoms of Environmental Structure', in MOORE, 1970.

APPS, P. (1976) 'Architectural Design and the Problem of Evaluation' in MARCH, 1976a.

ARCHER, L.B. (1964) Systematic Method for Designers, London, The Design Council/HMSO.

ARCHER, L.B. (1969) 'The Structure of the Design Process', in BROADBENT & WARD, 1969.

ARCHER, L.B. (1971) Technological Innovation - a Methodology, London, Science Policy Foundation.

ARCHER, L.B. (1972) 'Computers, Design Theory and the Handling of the Qualitative', paper to Computers in Architecture Conference, York.

ARCHER, L.B. (1974) Design Awareness and Planned Creativity in Industry, London, Design Council.

ARCHER, L.B. (1979) 'Design as a Discipline', Design Studies, 1, 1, 17-20.

ARENDT, H. (1958) The Human Condition, Chicago, University of Chicago Press.

ATKIN, R.H. (1976) 'A General Mathematical Approach in Physical, Urban and Social Systems', paper to IMA Symposium "The Mathematics of Urban Systems", University of Birmingham.

ATKINSON, G. (1978) 'Seeking a Yardstick for Quality', Building, 2 June, 47-50.

BAER, A.E. (1976) A Cognitive Model of Problem Formulation in Design, Research Report no.62, Institute of Physical Planning, Carnegie-Mellon University.

BAIER, K. (1973) 'The Concept of Value', in LASZLO & WILBUR, 1973.

BALDWIN, M.M. (ed) (1975) Portraits of Complexity, Battelle Monograph, 9, June 1975.

BANHAM, R. (1960) Theory and Design in the First Machine Age, London, Architectural Press.

BATTY, M. (1980) 'Limits to Prediction in Science and Design Science', Design Studies, 1, 3, 153-159.

BAYNES, K., LANGDON, R. & MYERS, B. (1977) 'Design in General Education: A Review of Developments in Britain', Art Education, 30, 8, 17-21.

BAYNES, K., LANGSLOW, B. & WADE, C.C. (1969) Evaluating New Hospital Buildings, London, King Edward's Hospital Fund.

BAZJANAC, V. (1974) 'Consequences of some Theories of Design on Computing in Architecture' in CARSON, 1974.

BEER, S. (1974) Designing Freedom, London, Wiley.

BENSASSON, S. & BURGESS, K. (1978) Computer Programs for Daylighting in Buildings, Evaluation Report no.3, Cambridge, Design Office Consortium.

BERGER, S.R. (1980) 'Artificial Intelligence and its Impact on Computer Aided Design', Design Studies, 1, 3, 166-171.

BESSANT, J.R. (1979) 'Preparing for Design Studies: ways of watching', Design Studies, 1, 2, 77-83.

BESSANT, J.R. & McMAHON, B.J. (1979) 'Participant Observation of a Major Design Decision in Industry', Design Studies, 1, 21-26.

BEST, A. (1978) 'Why The People Stay Away', Design, 354, 50-55.

BEVERIDGE, W.I. (1957) The Art of Scientific Investigation, London, Heinemann.

BIJL, A. (1975) Design: People and Computers, EdCAAD Report, University of Edinburgh.

BLOHM, H. & STEINBUCH, K. (1973) Technological Forecasting in Practice, Saxon House.

BOIES, S.V. & SPIEGEL, M.F. (1973) A Behavioual Analysis of Programming: on the use of interactive debugging facilities, New York, IBM Research Report.

BONO, E. de (1967) The Use of Lateral Thinking, New York.

BUILDING PERFORMANCE RESEARCH UNIT (1972) Building Performance, London, Applied Science.

BRITISH SHIP RESEARCH ASSOCIATION (1974) 'Detailed Study of Draughting Tasks using Closed Circuit Television', Report, BSRA, Wallsend.

BROADBENT, G.H. (1973) Design in Architecture, London, Wiley.

BROADBENT, G.H. (1977) 'Recent Developments in Design Method Studies', Paper to Workshop Zur Entwursforschung at Technische Hochschule, Aachen.

BROADBENT, G.H. (1979) 'Design and Theory Building', <u>Design Methods and Theories</u>, <u>13</u>, 3/4, 103-106.

BROADBENT, G.H. (1980) 'Architectural Objects and their Design as a Subject for Semiotic Studies', <u>Design Studies</u>, <u>1</u>, 4, 207-216.

BROADBENT, G.H. & WARD, A. (1969) <u>Design Methods in Architecture</u>, Architectural Association Paper no.4, London.

BROSS, I.D.J. (1953) <u>Design for Decision</u>, New York, Collier Macmillan.

BRUNER, J.S. (1960) <u>The Process of Education</u>, Cambridge, Mass., Harvard University Press.

BRUNER, J.S., GOODNOW, J.J. & AUSTIN, G.A. (1956) <u>A Study of Thinking</u>, New York, Wiley.

BRUTTON, M. (1978) 'Why Designers Should Stop Worrying And Learn To Love Design Evaluation', <u>Design</u>, <u>354</u>, p45.

BRYANT, D.A., DAINS, R.B. & SPOONAMORE, J.H. (1975) 'Structure of SEARCH-2', Report, Construction Engineering Research Laboratory, Champaign, Illinois.

BURT, M. (1978) <u>A Survey of Quality and Value in Building</u>, Garston, BRE.

CAKIN, S. (1975) 'Decision Making in Building Design: evaluation of holiday houses', <u>DMG-DRS Journal</u>, <u>9</u>, 2.

CAKIN, S. (1976) <u>Evaluation and Participation in Design</u>, PhD thesis, University of Strathclyde.

CAMPBELL, S. (1972) 'Architectural Values as a Measure of Design Decision Making', in MITCHELL, 1972.

CANTER, D.V. (1969a) <u>The Subjective Assessment of the Environment</u>, BPRU Report no.42, University of Strathclyde.

CANTER, D.V. (ed) (1969b) <u>Architectural Psychology</u>, London, RIBA.

CANTER, D. (1974) <u>Psychology for Architects</u>, London, Academic Press.

CANTER, D. (1977a) <u>The Psychology of Place</u>, London, Architectural Press.

CANTER, D. (1977b) 'Priorities in Building Evaluation: some methodological considerations', Journal of Architectural Research, 6, 38-40.

CANTER, D. & LEE, T. (eds) (1974) Psychology and the Built Environment, New York, Halsted.

CANTER, D. & WOOLS, R. (1970) 'A Technique for the Subjective Appraisal of Buildings', Building Science, 5, 3 & 4, 187-198.

CARROLL, V.M., THOMAS, J.C. & MALHOTRA, A. (1978a) A Clinical-Experimental Analysis of Design Problem Solving, IBM Research Report RC7081, Yorktown.

CARROLL, J.M., THOMAS, J.C. & MALHOTRA, A. (1978b) Presentation and Representation in Design Problem Solving, IBM Research Report RC6975, Yorktown.

CARSON, D.H. (1974) Man-Environment Interactions: Evaluations and Applications, Part III, Stroudsburg, Dowden, Hutchinson & Ross.

CHASE, W.G. & SIMON, H.A. (1973) 'Perception in Chess', Cognitive Psychology, 4, 55-81.

CHOMSKY, N. (1968) Language and Mind, New York, Harcourt, Brace & World.

CHURCHMAN, C.W. (1971) The Design of Inquiring Systems, basic concepts of systems and organisations, New York, Basic Books.

CHURCHMAN, C.W., ACKOFF, R.L. & ARNOFF, E.C. (1957) Introduction to Operations Research, New York, Wiley.

COLLINS, P. (1961) Changing Ideals in Modern Architecture, London, Faber and Faber.

COLLINS, P. (1971) Architectural Judgement, London, Faber and Faber.

CONWAY, D. (ed) (1974) Social Science and Design, Proceedings Coolfont Conference, Washington DC, AIA.

CONWAY, D. (ed) (undated) Architectural Design and the Social Sciences, Washington DC, AIA Research Report.

CORNFORD, C. (ed) (1977) 'Theoretical Studies and Art Education', RCA Papers, no.1.

CORNFORTH, C. (1976) Towards the Development of a Model of the Architectural Design Activity, MA thesis, DDR, RCA.

COURTIEUX, G. & GUIBERT, D. (1977) 'An Experiment in Teaching Computer Aided Design and Computer Aided Manufacturing in Architecture, Proceedings CADED Conference, Guildford, IPC.

CRONBERG, T. (1975) Performance Requirements for Buildings - a study based on user activities, Stockholm, Swedish Council for Building Research.

CROSS, N. (ed) (1972) Design Participation, London, Academy Editions.

CROSS, N. (1977) The Automated Architect, London, Pion.

CROSS, N. (1978) 'Assessing Computer-Aided Architectural Design Systems', Proceedings CAD 78 Conference, Guildford, IPC.

CROSS, N., ELLIOT, D. & ROY, R. (1975) Designing the Future, London, Open University Press.

DALEY, J. (1968) 'The Myth of Quantifiability', Architects' Journal, 148, 34, 339-341.

DALEY, J. (1969) 'A Philosophical Critique of Behaviourism in Architectural Design', in BROADBENT & WARD, 1969.

DANFORD, S. & WILLEMS, E.P. (1975) 'Subjective Responses to Architectural Displays: a question of validity', Environment and Behaviour, 7, 4, 486-516.

DARKE, J. (1979) 'The Primary Generator and the Design Process', Design Studies, 1, 1, 36-44.

DARWIN, R. (1959) Royal College of Art Annual Report, 1959, London, RCA.

DAVIS, C.F. & KENNEDY, M.D. (1970) 'EPS: A Computer Program for the Evaluation of Problem Structure', in MOORE, 1970.

DERBYSHIRE, M. (1976) 'Balancing Architectural Objectives - the relevance of decision theory', in MARCH, 1976a.

DRAKOU, C. (1973) Cognition and Design, MA thesis, School of Environmental Design, University College, London.

DUCKWORTH, E. (1962) A Guide to Operational Research, London, Methuen.

DUFFY, F. & TORREY, J. (1970) 'A Progress Report on the Pattern Language', in MOORE, 1970.

EASTMAN, C. (1970) 'On the Analysis of Intuitive Design Processes', in MOORE, 1970.

EASTMAN, C. (1975) The Concise Structuring of Geometric Data for Computer Aided Design, Research Report no.59, Institute of Physical Planning, Carnegie-Mellon University.

EASTMAN, C. (1980) 'Information and Databases in Design, Design Studies, 1, 3, 146-152.

EASTMAN, C. & HENRION, M. (1975) Language Features for a Design Information System, Report, Institute of Physical Planning, Carnegie-Mellon University.

EASTMAN, C. & SCHINNAR, A.P. (1974) A Stochastic Model of Performance for Adaptive Environments, Report, Institute of Physical Planning, Carnegie-Mellon University.

EDMUNDS, E. & LEE, J. (1974) 'An Appraisal of some Problems of Achieving Fluid Man/Machine Interaction', Proceedings Eurocomp Conference.

EDWARDS, M. (1974) 'Comparisons of Some Expectations of a Sample of Housing Architects with Known Data', in CANTER & LEE, 1974.

EDWARDS, W. (1967a) 'The Theory of Decision Making', in EDWARDS & TVERSKY, 1967.

EDWARDS, W. (1967b) 'Behavioural Decision Theory', in EDWARDS & TVERSKY, 1967.

EDWARDS, W. & TVERSKY, A. (eds) (1967) Decision Making, Harmondsworth, Penguin.

EMBREY, D. (1974) 'The Analysis of Design Activity: A Review and Redefinition of Problems, Aims and Techniques', Report 124/WP19, DDR, RCA.

EVANS, C.R., PURCELL, P.A. & WOOD, J. (1971) An Investigation of Design Activities Using Analytic Time-Lapse Photography, National Physics Laboratory Report no.Com Sci 50.

EVANS, P. (1973) 'Can Design be Taught?', Nature, 242, 441-2.

FENVES, S.J., RANKIN, K. & TEJUJA, H.K. (1976) The Structure of Building Specifications, Report, National Bureau of Standards Building Science Series 90, Washington.

FENVES, S.J. & WRIGHT, R.N. (1977) The Representation and Use of Design Specifications, National Bureau of Standards Technical Note 940, Washington.

FEREBEE, A. (1978) 'How a Brand New Metro Lost Its Way', Design, 354, 46-49.

FEYERABEND, P. (1975) Against Method, London, NLB.

FISHBURN, P.C. (1964) Decision and Value Theory, New York, Wiley.

FLORMAN, S.C. (1974) The Existential Pleasures of Engineering, New York (London, 1976 edition).

FOSTER, B.E. (ed) (1972) Performance Concept in Buildings, National Bureau of Standards, 2 vols., Washington.

FOZ, A.T.K. (1972) Some Observations on Designer Behaviour in the Parti, MA thesis, MIT.

FOZ, A.T.K. (1973) 'Observations on Designer Behaviour in the Parti', paper to Design Activity Conference, London.

FRIEDMAN, Y. (1975) Towards a Scientific Architect, Cambridge, Mass., MIT.

FRIEDMANN, A., ZIMRING, C. & ZUBE, E. (1978) Environmental Design Evaluation, New York, Plenum.

FRISCHMUTH, D.S. & ALLEN, T.J. (1969) 'A Model for the Description and Evaluation of Technical Problem Solving' IEEE Transactions Engineering Management, EM-16, 2, 58-64.

GASPARSKI, W. (1979a) 'Design Research - a tentative overview', Proceedings PArC 79 Conference, Berlin, AMK.

GASPARSKI, W. (1979b) 'Praxiological-Systematic Approach to Design Studies', Design Studies, 1, 2, 101-106.

GERO, J.S. (1980) 'Computer Aided Design by Optimization in Architecture', Design Studies, 1, 4, 227-230.

GIEDION, S. (1941) Space, Time and Architecture, Cambridge, Mass., Harvard University Press.

GILL, R. (1980) 'Approaches to Design', Design Studies, 1, 3, 141-145.

GOUMAIN, P. (1977) 'Towards Systems Tutors and Conversational Modes of Man-Machine Interaction', Report to DoE, DDR, RCA.

GOUMAIN, P. & MALLEN, G. (1978) 'Interface Design and the Future Development of Interactive Building Design Systems - a Study Report', Proceedings CAD 78 Conference, Guildford, IPC.

GRANT, D.P. (1974) 'The Problem of Weighting', DMG/DRS Journal, 8, 3, 136-141.

GRANT, D.P. (1976) 'The "How To" of Decision Making', Design Methods and Theories, 10, 4, 200-237.

GRANT, D.P. (1978) 'Response to James Wise's Critical Reply to the Alpha-Beta Model for Decision Making', Design Methods and Theories, 12, 1, 44-45.

GREEN, P.E., CARMONE, F.J. & WIND, Y. (1972) 'Subjective Evaluation Models and Conjoint Measurement', Behavioural Science, 17, 288 - 299.

GREGORY, S.A. (1979a) 'New Materials Adoption Study: some contributions to design knowledge', Design Studies, 1, 2, 107 - 112.

GREGORY, S.A. (1979b) 'What We Know About Designing And How We Know It', paper to Institute of Chamical Engineering Design Congress, University of Aston.

GROPIUS, W. (1935) The New Architecture and the Bauhaus, London, Faber & Faber. (1955 edition).

GUTMAN, R. & WESTERGARD, B. (1978) 'What Architectural Schools Know about their Graduates', Journal of Architectural Education, 31, 2, 2 - 11.

HARARY, F. (1964) 'A Graph Theoretic Approach to Similarity Relations', Psychometrika, 29, 2, 143 - 151.

HARPER, D.H. (1978) Building: The Process and the Product, Hornby, The Construction Press.

HARTMAN, R.S. (1973) 'Formal Axiology and the Measurement of Values', in LASZLO & WILBUR, 1972.

HAWKES, D. & STIBBS, R. (1969) The Environmental Evaluation of Buildings: 1 A mathematical model, LUBFS Working paper no.15, Cambridge.

HAWKES, D. & STIBBS, R. (1970a) The Environmental Evaluation of Buildings: 2 Technical specifications of a model, LUBFS Working paper no.27, Cambridge.

HAWKES, D. & STIBBS, R. (1970b) The Environmental Evaluation of Buildings: 3 A worked example, LUBFS Working paper no.28, Cambridge.

HAWKES, D. & STIBBS, R. (1970c) The Environmental Evaluation of Buildings: 4 The description and evaluation of rooflights, LUBFS Working paper no.29, Cambridge.

HEATH, T.F. (1968) 'Problems of Measurement in Environmental Aesthetics', Architectural Science Review, March 1968, 17 - 28.

HEATH, T.F. (1970) 'The Algorithmic Nature of the Design Process', in MOORE, 1970.

HENRION, M. (1974) Notes on the Synthesis of Problems, MA thesis, DDR, RCA.

HESSELGREN, S. (1967) The Language of Architecture, Sweden, Student-litteratur (and London, 1972, Applied Science).

HESSELGREN, S. (1975) Man's Perception of Man-Made Environment, Lund, Student-litteratur.

HILLIER, W.R.G. (1969) 'Psychology and the Subject Matter of Architectural Research', in CANTER, 1969.

HILLIER, B., LEAMAN, A., STANSALL, P. & BEDFORD, M. (1976) 'Space Syntax', Environment and Planning B, 3, 147-185.

HILLIER, B., MUSGROVE, J. & O'SULLIVAN, P. (1972) 'Knowledge and Design', in MITCHELL, 1972.

HONIKMAN, B. (ed) Proceedings AP 70 Conference, London, RIBA.

HONIKMAN, B. (1973) 'Personal Construct Theory and Environmental Evaluation', in PREISER, 1973.

HOLLAND, V. (1972) 'Cognitivie Persuations: assumptions and presumptions', in MITCHELL, 1972.

HONIKMAN, B. (1976) 'Environmental Evaluation', Journal of Architectural Education, XXIX, 4, 14-15.

HOPKINSON, R.G. (1966) The Evaluation of the Built Environment, London, H.K. Lewis.

HORMANN, A.M. (1972) 'Machine-Aided Evaluation of Alternative Designs', in MITCHELL, 1972.

HUBER, G.P., SAHNEY, V. & FORD, D. (1969) 'A Study of Subjective Evaluation Models', Behavioural Science, 14, 483-489.

HUBERT, L.J. (1974) 'Some Applications of Graph Theory to Clustering', Psychometrika, 39, 283-309.

JENCKS, C. (1973) Modern Movements in Architecture, Harmondsworth, Penguin.

JOHNSON, J.P. (1973) 'The Fact-Value Question in Early Modern Value Theory', in LASZLO & WILBUR, 1973.

JOHNSON, S.C. (1967) 'Hierarchical Clustering Schemes', Psychometrika, 32, 3, 241-255.

JONES, J.C. (1970) Design Methods: Seeds of Human Futures, London, Wiley.

JONES, J.C. & THORNLEY, D.G. (eds) (1963) Conference on Design Methods, Oxford, Pergamon.

KAPLAN, B. (1970) 'On the Rational Reconstruction of Intuition in the Design Process', in MOORE, 1970.

KAPLAN, R. (1972) 'The Dimensions of the Visual Environment: methodology considerations', in MITCHELL, 1972.

KAPLAN, R. (1973) 'Predicters of Environmental Preference: Designers and Clients', in PREISER, 1973.

KAUFMAN, G.M. & THOMAS, H. (1977) Modern Decision Analysis, Harmondsworth, Penguin.

KEENY, R.L. & RAIFFA, H. (1976) Decisions with Multiple Objectives, New York, Wiley.

KOHLER, W. (1938) The Place of Value in a World of Facts, New York, Liveright.

KRAUSS, R.I. & MYER, J.R. (1970) 'Design: a case history', in MOORE, 1970.

KUHN, T.S. (1962) The Sturcture of Scientific Revolutions. Chicago, University of Chicago Press.

LASZLO, E. & WILBUR, J.B. (eds) (1973) Value Theory in Philosophy and Social Science, New York, Gordon & Breach.

LAWSON, B.R. (1970) 'Open and Closed Ended Problem Solving in Architectural Design', in HONIKMAN, 1970.

LAWSON, B.R. (1972) Problem Solving in Architectural Design, PhD thesis, University of Aston.

LAWSON, B.R. (1978) 'The Architect as a Designer' in SINGLETON, 1978.

LAWSON, B.R. (1979) 'The Act of Designing', Design Methods and Theories, 13, 1, 29-32.

LEFT, H.S. & DEUTSCH, P.S. (1973) 'Construing the Physical Environment: Differences between Environmental professionals and lay persons', in PREISER, 1973.

LERA, S.G. (1976) The Architecture of Design, BArch dissertation, University of Liverpool.

LERA, S.G. & THOMPSON, B.G.J. (1980) 'Computer Aided Design in Practice', RIBA Journal, 87, 5, 67-69.

LEVIN, P. (1966) Decision Making in Urban Design, BRE Current Papers, Design Series no.49.

LOWE, J.B. (1970) 'The Assessment of Students' Architectural Design Drawings', Architectural Research & Teaching, 1, 37-45.

LOWE, J.B. (1972) 'The Assessment of Students' Architectural Design Drawings', Architectural Research & Teaching, 2, 96-104.

LUCE, R.D. & RAIFFA, H. (1957) Games and Decisions, New York, Wiley.

MACKINNON, D.W. (1970) 'The Personality Correlates of Creativity: A study of American architects', in VERNON, 1970.

MAGEE, B. (1973) Popper, London, Fontana.

MAGEE, B. (ed) (1978) Men of Ideas, London, BBC.

MALLEN, G.L. (1973) 'The Role of Simulation in Social Education', Programmed Learning and Educational Technology, 10, 4, 248 - 258.

MALLEN, G.L. (1978) 'Project DoE/MCI: outline of visit to USA', Report, Systems Simulation Ltd., London.

MALLEN, G.L. & GOUMAIN, P. (1973) The Analysis of Architectural Design Activity in the Working Environment, Report 108/3, DDR, RCA.

MALLEN, G.L. & GOUMAIN, P. (1976) 'A Philosophy for Research (a Framework for design research), Seminar papers, DDR, RCA.

MALLEN G.L., GOUMAIN, P.G.R. & PURCELL, P.A. (1974) A Strategy for Design Research, Report 108/4, DDR, RCA.

MARCH, J.G. & SIMON, H.A. (1958) Organisations, New York, Wiley.

MARCH, L. (1976a) The Architecture of Form, Cambridge, Cambridge University Press.

MARCH, L. (1976b) 'The Logic of Design and the Question of Value', introduction to MARCH, 1976a.

MARGOLIS, J. (1973) 'The Use and Syntax of Value Judgements' in LASZLO & WILBUR, 1973.

MARKUS, T.A. (1969) 'The Role of Building Performance, Measurement and Appraisal in Design Method', in BROADBENT & WARD, 1969.

MAVER, T.W. (1970) 'Appraisal in the Building Design Process', in MOORE, 1970.

MAVER, T.W. (1975) 'Three Design Paradigns: a tentative philosophy', DMG/DRS Journal. 9, 2, 130 - 132.

MAVER, T.W. (1976) 'Democracy in Design Decision-Making', Proceedings CAD 76 Conference, Guildford, IPC.

MAVER, T.W., SMITH, M., WATTS, J. & AISH, R. (1979) 'Implications for Practice and Education', Proceedings PArC 79 Conference, Berlin, AMK.

MAVER, T.W. (1980) 'Appraisal in Design', Design Studies, 1, 3, 160 - 165.

MAYALL, W.H. (1979) Principles in Design, London, Design Council.

MEDAWAR, P.B. (1969) Induction and Intuition in Scientific Thought, London, Methuen.

MICHON, J.A. (1969) 'Multi-Dimensional and Hierarchical Analysis of Progress in Learning', paper to Fifth Annual Symposium on Cognition, Carnegie-Mellon University.

MILLER, D.W. & STARR, M.K. (1967) The Structure of Human Decisions, Englewood Cliffs, N.J., Prentice Hall.

MILLER, G.A. (1956) 'The Magical Number Seven, Plus or Minus Two', Psychological Review, 63, 2, 81-97.

MILLER, G.A., GALANTER, E. & PRIBRAM, K.H. (1960) Plans and the Structure of Behaviour, New York, Holt, Rinehart & Winston.

MITCHELL, W. (ed) (1972) Environmental Design Research and Practice, Los Angeles, EDRA.

MOORE, G.T. (1970) Emerging Methods in Environmental Design and Planning, Cambridge, Mass., MIT.

MORAN, T. (1970) 'A Model of a Multilingual Designer', in MOORE, 1970.

MORONEY, M.J. (1951) Facts from Figures, Harmondsworth, Penguin.

MURDOCH, I. (1970) The Sovereignty of Good, London, Routledge & Kegan Paul.

MUSGROVE, J. (1975) 'A Brief History of Architectural Research in the Past Three Decades', paper to Architectural Research Day, RIBA.

NEGROPONTE, N. (1976) 'Sketching: A Computational Paradigm for Personalised Searching', Journal of Architectural Education, XXIX, 2, 26-29.

NEUFVILLE, R. de & KEENEY, R.L. (1977) 'Systems Evaluation through Decision Analysis: Mexico City Airport', in KAUFMAN & THOMAS, 1977.

NEUMANN, J. von & MORGENSTERN, O. (1947) Theory of Games and Economic Behaviour, Princeton, Princeton University Press.

NEWELL, A. & SIMON, H.A. (1972) Human Problem Solving, Englewood Cliffs, N.J., Prentice Hall.

NIXON, D, PERERA, N. & GOUMAIN, P. (1979) Multi-purpose Industrial Development, Report 201, DDR, RCA.

NORRIS, K.W. (1963) 'The Morphological Approach to Engineering Design', in JONES & THORNLEY, 1963.

NOWELL-SMITH, P.H. (1954) Ethics, Harmondsworth, Penguin.

O'KEEFE, J.P. (1974) 'Computer Aided Design as a Teaching Aid in Architectural Education', International Journal of Mathematical Education in Science and Technology, 5, 3, 375-379.

OSBORN, A.F. (1957) Applied Imagination, New York, Scribners.

O'SULLIVAN, P. & TERRITT, C. (1975) 'Selection and Sponsorship of University Research', paper to Architectural Research Day, RIBA.

OZBEKHAN, H. (1971) 'Planning and Human Action', in WEISS, 1971.

PAGE, J.K. (1963) 'A Review of the Papers Presented at the Conference' in JONES & THORNLEY, 1963.

PETROVIC, I.K. (1972) System and Design, PhD thesis, Department of Architecture, Queen's University, Belfast.

PHILLIPS, L.D. (1978) 'Approaches to Decision Making', tutorial paper 78-1, Brunnel University.

PIRSIG, R.M. (1974) Zen and the Art of Motorcycle Maintenance, London, Bodley Head.

PORAT, A.M. & HAAS, J.A. (1969) 'Information Effects on Decision Making', Behavioural Science, 14, 98-104.

PREISER, W.F.E. (ed) (1973) Environmental Design Research, Stroudsberg, Dowden Hutchinson & Ross.

PROSHANSKY, H.M., ITTELSON, W.H. & RIVLIN, L.G. (eds) (1970) Environmental Psychology: man and his physical setting, New York, Holt Rinehart & Winston.

RADFORD, A.D. & GERO, J.S. (1979) 'Optimization for Information in Integrated Environmental Design', Proceeding PArC 79 Conference, Berlin, AMK.

RAIFFA, H. (1968) Decision Analysis, Reading, Mass., Addison-Wesley.

RAMAN, P.G. (1973) 'Synthesis in Design - An Interdisciplinary Essay', in PREISER, 1973.

REITMAN, W.R. (1964) 'Heuristic Decision Procedures, Open Constraints, and the Structure of Ill-Defined Problems', in SHELLY & BRYAN, 1964.

RESCHER, N. (1969) Introducation to Value Theory, Englewood Cliffs, N.J., Prentice Hall.

RESCHER, N. (1973) 'The Study of Value Change', in LASZLO & WILBUR, 1973.

RIBA (1967) Handbook of Architectural Practice and Management, London, RIBA.

RIBA INTELLIGENCE UNIT (1975) 'Defing Architectural Research', paper to Architectural Research Day, RIBA.

RIBA INTELLIGENCE UNIT (1975) 'Current Problems in Building Design Research', paper to Architectural Research Day, RIBA.

RICKABY, P. (1979) 'Speculations on the Future Practice of Architecture', Design Studies. 1, 2, 118-123.

ROBERTSON, W.K. (1978) A Contribution Towards a Methodology by which Design Skill may be Studied, MA thesis, DDR, RCA.

ROKEACH, M. (1973) The Nature of Human Values, London, Collier Macmillan.

RUSCH, C.W. (1970) 'The Role of Graphic Activity in the Design Process', in MOORE, 1970.

RUSKIN, D.J. (1978) 'Weighting of Partial Judgements in Pbjectification and Evaluation of Decisions', Design Methods and Theories. 12, 1, 32-39.

SAATY, T.L. (1977) 'A Scaling Method for Priorities in Hierarchical Structures', Journal of Mathematical Psychology. 15, 3, 234-281.

SAATY, T.L. (1978) 'Modelling Unstructured Decision Problems - The Theory of Analytical Hierarchies', Mathematics and Computers in Simulation, XX, 147-158.

SAATY, T.L. (1979) 'Applications of Analytical Hierarchies', Mathematics and Computers in Simulation, XXI, 1-20.

SAATY, T.L. & ERDENER, E. (1979) 'A New Approach to Performance Measurement: the Analytic Hierarchy Process', Design Methods and Theories, 13, 2, 64-72.

SAATY, T.L. & ROGERS, P.C. (1976) 'Higher Education in the United States 1985-2000', Socio-Econ. Plan. Sci., 10, 251-263.

SAVAGE, L.J. (1967) 'Historical and Critical Comments on Utility', in EDWARDS & TVERSKY, 1967.

SHELLY, M.W. & BRYAN, G. (eds) (1964) Human Judgements and Optimality, New York, Wiley.

SHEPARD, R.N. (1967) 'On Subjectively Optimum Selections among Multi-attribute Alternatives', in EDWARDS & TVERSKY, 1967.

SIDALL, J.N. (1972) 'Value Theory as a Vehicle for User Participation', in CROSS, 1972.

SIEGEL, S. (1956) Non Parametric Statistics for the Behavioural Sciences, New York, McGraw-Hill.

SIMON, H.A. (1960) The New Science of Management Decision, Englewood Cliffs, N.J., Prentice Hall (1977 ed).

SIMON, H.A. (1969) The Sciences of the Artificial, Cambridge, Mass., MIT.

SIMON, H.A. (1973) 'The Structure of Ill-Structured Problems', Artificial Intelligence, 4, 181 - 201.

SIMPSON, O.von (1956) The Gothic Cathedral, London, Routledge & Kegan Paul.

SINGLETON, W.T. (ed) (1978) The Analysis of Practical Skills, Lancaster, MTP.

STANSALL, P.J. (1973) A Contribution toward the Development of Models of Thought-Processes underlying the Design Activity, MA thesis, DDR, RCA.

STEADMAN, P. (1979) The Evolution of Designs, Cambridge, Cambridge University Press.

STRINGER, P. (1969) 'Architecture, Psychology the Game's the Same', in CANTER, 1969b.

TAGG, S.K. (1974) 'The Subjective Meaning of Rooms: some analyses and investigations', in CANTER & LEE, 1974.

TARNOWSKI, W. (1979) 'General Model and Choice Criterion in Engineering Design', Design Studies, 1, 1, 45 - 48.

THOMAS, J.C. & CARROLL, J.M. (1979) 'The Psychological Study of Design', Design Studies, 1, 1, 5 - 11.

THOMPSON, B.G.J. (1979) 'Guildford Crown Courts sketch design appraisal by CEDAR3', CAD Journal, 11, 5, 305 - 307.

THOMPSON, B.G.J. & HUGHES, J. (1974) 'An Experimental Investigation of the Performance of a Computer-Aided Building Design System', Proceedings IFIP 74 Conference, Stockholm.

THOMPSON, B.G.J., LERA, S.G., BEESTON, D. & COLDWELL, R. (1978) 'Pre-production Trials of CEDAR3', Internal Report, PSA DoE.

THOMPSON, B.G.J., LERA, S.G., BEESTON, D. & COLDWELL, R. (1979) 'Application of CEDAR3: case studies from the pre-production trials of a large interactive computer aided building design system in United Kingdom Central Government organisations', *Proceedings PArC 79 Conference*, Berlin, AMK.

THOMPSON, B.G.J. & WEBSTER, G.J. (1978) 'Progress with CEDAR3: a computer aided building design system for the sketch plan stage', *Proceedings CAD 78 Conference*, Guildford, IPC.

THOMPSON, B.G.J. & WOOD, J. (1975) *An Intensive Design Exercise to study the use of the PSA Method of Building with a view to possible requirements for computer aids*, DoE PSA Report, Croydon.

THOMPSON, B.G.J. & YOUNG, J.S. (1980) 'Using CEDAR3 in Professional Practice', *Proceedings CAD 80 Conference*, Guildford, IPC.

THORNLEY, D.G. (1963) 'Design Method in Architectural Education', in JONES & THORNLEY, 1963.

TRIBE, L.H., SCHELLING, C.S. & VOSS, J. (eds) (1976) *When Values Conflict*, Cambridge, Mass., Ballinger.

TRIBUS, M. (1969) *Rational Descriptions, Decisions and Designs*, New York, Pergamon.

TSCHUDI, F. (1972) *The Latent, the Manifest and the Reconstructed in Multi variate Data Reduction Models*, PhD thesis, University of Oslo.

TVERSKY, A. (1967) 'Additivity, Utility and Subjective Probability', in EDWARDS & TVERSKY, 1967.

TVERSKY, A. & KAHNEMAN, D. (1977) 'Judgements under Uncertainty: heuristics and biases', in KAUFMAN & THOMAS, 1977.

TZONIS, A. & SALAMA, O. (1975) 'Problems of Judgements in Programmatic Analysis in Architecture: the synthesis of partial evaluations', *Journal of Architectural Research*, *4*, 47-54.

VERNON, G.M. (1973) 'Values, Value Definitions and Symbolic Interaction', in LASZLO & WILBUR, 1973.

VERNON, P.E. (1970) *Creativity*, Harmondsworth, Penguin.

VICKERS, G. (1968) *Value Systems and Social Process*, New York, Basic Books.

WADE, J.W. (1977) *Architecture, Problems, and Purposes: architectural design as a basic problem solving process*, New York, Wiley.

WATKIN, D. (1977) *Morality and Architecture*, Oxford, Clarendon.

WEINBERG, G.M. (1972) *The Psychology of Computer Programming*, New York, Van Nostrand.

WEINZAPFEL, G. (1973) 'It Might Work, But Will It Help?', paper to Design Activity Conference, London.

WEISS, P.A. (ed) (1971) *Hierarchically Organised Systems*, New York, Hafner.

WIENER, N. (1948) *Cybernetics. Control and Communication in the Animal and the Machine*, New York, Wiley.

WILCOX, J.W. (1972) *A Method for Measuring Decision Assumptions*, Cambridge, Mass., MIT.

WILLEY, D.S. & YEOMANS, D.T. (1974) 'Monitoring Graphic Techniques in Design', paper to CAD Conference, ICL, September 1974.

WILLS, G. & WILBUR, R. (1972) *Technological Forecasting*, Harmondsworth, Penguin.

WISE, J.A. (1978) 'A Critical Reply to Donald Grant's "Alpha-Beta" Model for Decision Making with Multiple Objectives', *Design Methods & Theories*, 12, 1, 40-43.

WITTKOWER, R. (1962) *Architectural Principlies in the Age of Humanism*, London, Tiranti.

WOOLS, R. (1969) 'The Assessment of Room Friendliness' in CANTER, 1969b.

YNTEMA, D.B. & TORGERSON, W.S. (1967) 'Man-Computer Cooperation in Decisions Requiring Common Sense', in EDWARDS & TVERSKY, 1967.