5 Measurement, criteria and judgement in design

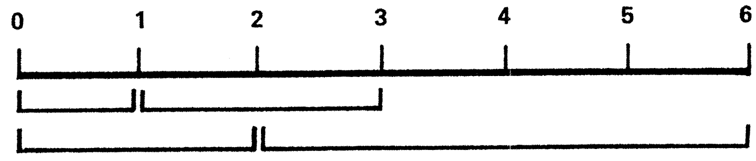
"She can't do Substraction," said the White Queen. "Can you do Division? Divide a loaf by a knife – what's the answer to that?" "I suppose –" Alice was beginning, but the Red Queen answered for her. "Bread-and-butter of course."

Lewis Carroll, *Alice Through the Looking Glass*

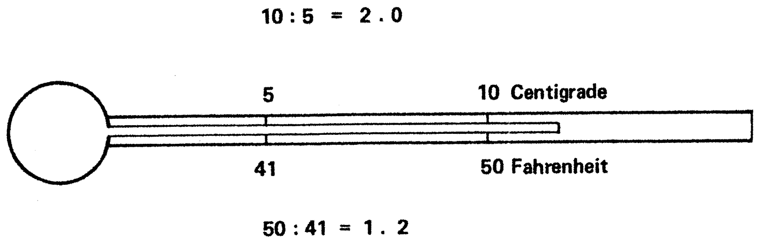
Measurement

Learning and developing a good design process is not an end in itself, but hopefully just a step towards producing better designs. In the final analysis it is the solution not the process which matters in design. But how can the degree of success achieved by a design solution be measured? Is it possible to say that one design is better than another and if so by how much? Such questions are not so easily answered as it may appear on first examination.

Consider the design of a garden greenhouse. There are a number of features which the designer of a greenhouse can vary. He could choose between several different materials for the frame; perhaps wood, steel, aluminium or plastic. The actual form of the greenhouse is even more variable with possibilities of domes, tent shapes, barrel vaults and so on. In fact there are many more design variables including the glazing material, method of ventilation and type of door. What the designer has to do is to select the combination of all these features which will give the most satisfactory performance. How then do we measure the performance of our greenhouse? A greenhouse is designed for one fairly clear and simple purpose; to trap heat from the sun, so we can begin by measuring or calculating the thermal efficiency of a whole range of possible greenhouses. Unfortunately, we are still some way from describing how satisfactory our greenhouse will appear to a gardener. He may well also want to know how much it will cost to buy, how long it will last, or how easy it will be to erect and maintain, and perhaps, what it will look like in his garden. The greenhouse then, must



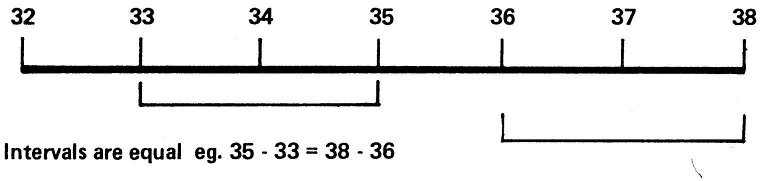
5.1 The ratio scale of measurement



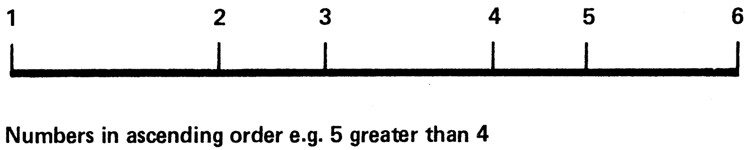
5.2 On thermometers, temperature is obviously not a ratio scale since the ratios of two temperatures are different in centigrade and fahrenheit

satisfy criteria of solar gain, cost, durability, ease of assembly, appearance and perhaps many others. It is quite likely that these criteria are not all equally important but the real difficulty here is that neither are they easily related one to another. It is obviously relatively easy to measure solar gain or durability, but what about ease of assembly or appearance? Measurement in design apparently involves both quantities and qualities. Somehow then, designers must be able to balance both qualitative and quantitative criteria in their decision making process.

In every day life we tend to use numbers and scales of measurement very carelessly. In fact we commonly employ several quite distinct ways of using numbers, without really being aware of the differences. Designers cannot afford to be so careless since they must frequently use all these different kinds of measurement scales simultaneously. We normally tend to assume that numbers are organised along what is usually called a ratio scale which is the normal language of counting where four represents twice two. However this is not always the case, for example while four apples certainly are twice as many as two apples, ten degrees centigrade are not twice as hot as five degrees centigrade. Why this should be can be seen by using both our common temperature scales together. One temperature described as 10° centigrade can also



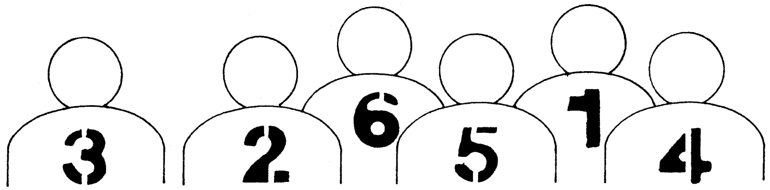
5.3 The interval scale of measurement



5.4 The ordinal scale of measurement

be described as 50° fahrenheit, and a lower temperature of 5° centigrade corresponds to 41° fahrenheit. Thus these two temperatures give a ratio of 10 to 5 or 2 to 1 on the centigrade scale, but a ratio of 50 to 41 on the fahrenheit scale. This is because the zero point on these scales is not absolute but entirely arbitrary. The centigrade scale is actually defined as having one hundred equal intervals between the freezing and boiling temperatures of water. We could equally easily use a thousand intervals or the freezing and boiling temperatures of any other substance. These temperature scales are described as interval measurement. Although 10°C cannot be described as twice as hot as 5°C the difference, or interval, between 0°C and 5°C is exactly equal to the interval between 5°C and 10°C. Interval scales are frequently used for subjective assessment. Psychologists recommend that such scales should be fairly short, up to seven intervals, to retain the reliability of the interval. Thus to return to our greenhouse, we might ask a number of gardeners to assess the ease of assembly or maintenance on five point scales. We would not be justified in regarding a greenhouse assessed as 4 for assembly as being twice as easy to assemble as one assessed as only 2.

Sometimes we use an even more cautious scale of measurement where the interval is not considered to be reliably consistent. Such scales are called ordinal, for they represent only a sequence or order. Cars in a traffic queue are arranged along an ordinal scale. We make no guarantee of the size of gap between each car, when we describe a particular car as fourth or fifth in the queue. The winners of a race are



5.5 The nominal scale of measurement

described as first, second and third, but this does not tell us how large were the gaps between them. Regulations require that the materials used in buildings should not allow flame to spread across their surface in case of fire. Materials can belong to one of five surface spread of flame classes which range from class 0 to class 4. On this ordinal scale the higher the number the more rapidly flame will spread, but the difference between class 1 and class 2 is not necessarily the same as the difference between class 2 and class 3. We also get ordinal scales when we ask people to rank order their preferences. Thus we could ask our gardeners to put a number of greenhouses in order of attractiveness of appearance. Whether ordinal or interval scales of assessment are appropriate remains a matter of judgement, but generally ordinal scales should be used where the assessment may depend on many factors or where the factors cannot easily be defined. Thus while it seems reasonable to ask our gardeners how much easier it is to assemble one greenhouse than another, it does not seem reasonable to ask how much more attractive it may be. Academic examiners may award marks out of one hundred for a particular examination, which is really an interval scale since the zero point is rarely used. Overall degree classifications however are usually based on the cruder ordinal scale of first, upper and lower second, third and pass.

Finally the fourth, least precise numbering system in common use is the nominal scale, so called because the numbers really represent names and cannot be manipulated arithmetically. The numbers on football players' shirts are nominal. A forward is neither better nor worse than a back and two goal keepers do not make a full back. In fact there is no sequence or order to these numbers, we could equally easily have used the letters of the alphabet or any other set of symbols.

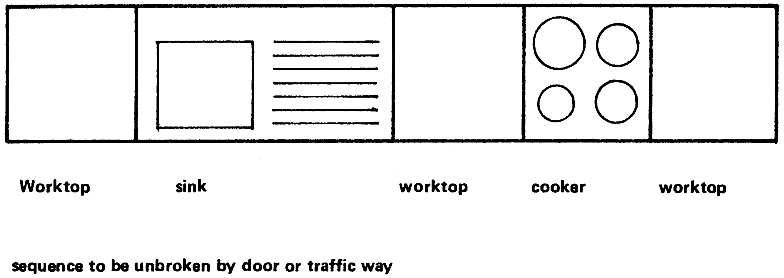
The importance of understanding these scales of measurement lies in the recognition that they must each be used in different ways. The inappropriate use of too precise a scale or arithmetic may lead to misleading results and false conclusions in design. One of the most well-known cases of such a confusion between scales of measurement is to be found in Archer's (1969) highly elaborate and numerical model of the design process. Archer, apparently somewhat reluctantly, concedes that at least some assessment of design must be subjective, but since he sets up a highly organised system of measuring satisfaction in design Archer clearly wants to use only ratio scales. Archer argues that a scale of 1–100 can be used for subjective assessment and the data then treated as if it were on a true ratio scale. In this system a judge, or arbiter as Archer calls him, is asked not to rank order or even to use a short interval scale, but to award marks out of 100. Archer argues that if the arbiters are correctly chosen and the conditions for judgement are adequately controlled, such a scale could be assumed to have an absolute zero and constant intervals. Archer does not specify how to "correctly choose" the judges or "adequately control the conditions," and his argument does seem rather suspect. In fact Stevens (1951), who originally defined the rules for measurement scales, did so to discourage psychologists from exactly this kind of numerical dishonesty. It is interesting to note that psychology itself was then under attack in an age of logic as being too imprecise to deserve the title of science. Perhaps for this reason, many psychologists had been tempted to treat their data as if it were more precise than Stevens' rules would indicate. Archer's work seems a parallel attempt to force design into a scientifically respectable mould. It now seems, some years later, that this attempt has failed and it is important to understand just why Stevens' rules for measurement scales should be respected in design as much as in psychology.

Value judgement and criteria

It is frequently tempting to employ more apparently accurate methods of measurement in design than the situation really deserves. Not only do the higher level scales, ratio and interval, permit much more arithmetic manipulation, but they also permit absolute judgement to be made. If it can be shown that, under certain circumstances 20°C is found to be a comfortable temperature then that value can be used as an absolutely measurable criterion of acceptability. Life is not so easy when ordinal measurement must be used. Universities use external examiners to help protect and preserve the "absolute" value of their degree classifications. It is, perhaps, not too difficult for an experienced examiner to rank order his pupils, but much more difficult to maintain a constant standard over many years of developing curricula and changing examinations. It is tempting to avoid these difficult problems of

judgement by instituting standardised procedures. Thus, to continue the example, a computer-marked multiple choice question examination technique might be seen as a step towards more reliable assessment. But there are invariably disadvantages with such techniques. Paradoxically, conventional examinations allow examiners to tell much more accurately, if not entirely reliably, how much their students have actually understood.

It is easy to fall into the trap of over-precision in design. Students of architecture sometimes submit thermal analyses of their buildings with the rate of heat loss through the building fabric calculated down to the last watt. Ask them how many kilowatts are lost when a door is opened and they are incapable of answering. Daylight design has been the subject of perhaps the most absurd series of measurement misconceptions. We all know that the level of natural illumination varies considerably throughout the day and even from minute to minute if the sun is periodically obscured by clouds. Normally this is relatively unimportant since the human eye is able to accommodate a very wide range of illumination. In fact the ratio of the brightest to the dimmest levels of illumination at which the human eye works efficiently is about 100,000 : 1. However this appears to be too untidy for building scientists who have invented the concept of a standard overcast sky. In this notional sky brightness is uniform all round the compass but greater at the zenith than the horizon, and there is no awkward sun. Such an abstraction permits the calculation of the percentage of total illumination from this theoretical sky which is available at a particular point in a room. Because these calculations involve very complex solid geometry the designer can easily be misled into believing that he has arrived at an absolute and meaningful result. Nothing could be further from the truth. In fact the daylight factor can only be used to compare one point with another on paper, it tells very little about the actual experience of being in the room in question. The danger of such techniques is that sooner or later they get used as fixed criteria. In fact this actually happened in the case of daylighting. Using statistics of the actual levels of illumination expected over the year it was calculated that a 2 per cent daylight factor was desirable in schools. It then became a mandatory requirement that all desks in new schools receive at least this daylight factor. The whole geometry of the classrooms themselves was thus effectively prescribed and, as a result, a generation of schools were built with large areas of glazing. The resultant acoustic and visual distraction, glare, draughts, the colossal heat losses and excessive solar gain in summer, which were frequently experienced in these schools,



5.6 A mandatory design requirement does not allow the designers to make a value judgement

eventually led to the relaxation of this regulation.

Unfortunately, much of the legislation with which designers must work appears to be based on the pattern illustrated by the daylighting example. Wherever there is the possibility of measuring performance, there is also the opportunity to legislate. It is difficult to legislate for qualities, but easy to define and enforce quantities. It is increasingly difficult for the designer to maintain a sensibly balanced design process in the face of necessarily imbalanced legislation. One of the most heavily legislated design problems today is to be found in public sector housing. Apart from the building regulations, some of which are especially demanding in the case of housing, the architect has to observe the government mandatory minimum standards for local authority housing. Many of these requirements were drawn out of the 1961 Parker Morris research report on "Homes for Today and Tomorrow". It is interesting to see which of the research committees many recommendations actually became mandatory requirements and which did not. Consider three of the original Parker Morris recommendations about kitchens.

1. The relation of the kitchen to the place outside the kitchen where the children are likely to play should be considered.
2. A person working at the sink should be able to see out of the window.
3. Worktops should be provided on both sides of the sink and cooker positions. Kitchen fitments should be arranged to form a work sequence comprising worktop/sink/worktop/cooker/worktop unbroken by a door or any other traffic way.

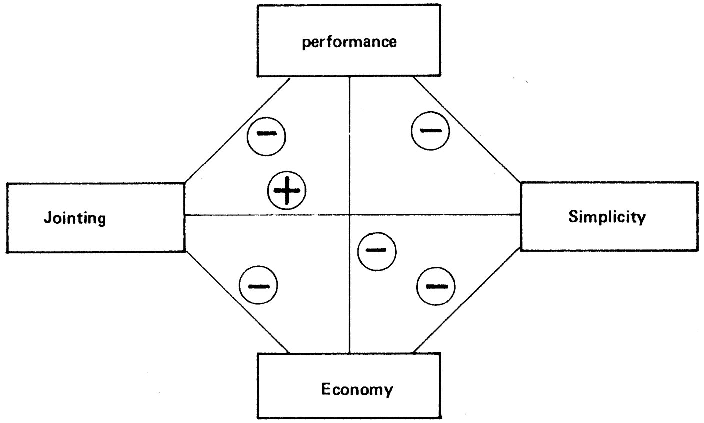
All these recommendations seem sensible and desirable. As criteria of performance however they are not all so easily measured from an architect's drawing. In fact only the last recommendation became a mandatory requirement. Thus it is now quite permissible to design a

family maisonette or flat off the ground with no view of any outside play spaces from the kitchen, but with the very model of a kitchen work surface as may not be found even in some very expensive privately built housing. It is worth noting that this legislation was introduced during the early period of what has now been called first generation design methodology. Reference has already been made to Alexander's (1964) famous method of design, which perhaps exemplifies the thinking about design at this time, and we shall pause here to fill in some detail.

Alexander's method involved first listing all the requirements of a particular design problem, and then looking for interactions between these requirements. For example in the design of a kettle some requirements for the choice of materials might be as follows.

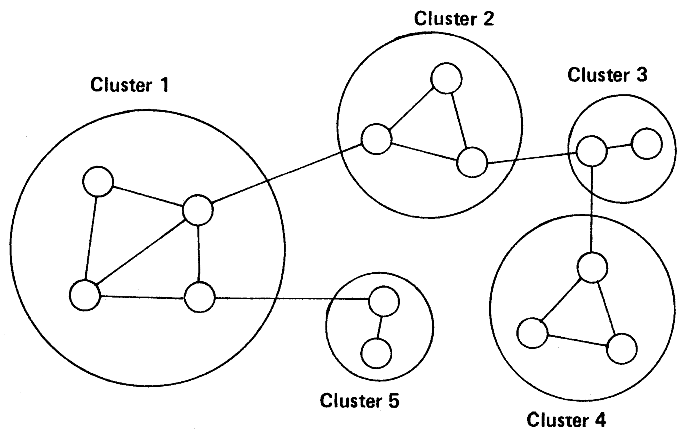
Simplicity:	the fewer the materials the more efficient the factory
Performance:	each function within the kettle requires its own material e.g. handle, lid, spout
Jointing:	the fewer the materials the less and the simpler the jointing and the less the maintenance
Economy:	choose the cheapest material suitable

The interactions between each pair of these requirements are next labelled as positive, negative or neutral depending on whether they complement, inhibit or have no effect upon each other. In this case all the interactions except jointing/simplicity are negative since they show conflicting requirements. For example while the performance requirement suggests many materials, the jointing and simplicity



5.7 The requirements and their interactions for "Alexander's kettle" requirements would ideally be satisfied by using only one material. Thus jointing and simplicity interact positively with each other but both interact negatively with performance. Thus a designer using Alexander's method would first list all the requirements of his design and then state which pairs of requirements interact either positively or negatively. All this data would then be fed into a computer program which looks for clusters of requirements which are heavily inter-related but relatively unconnected with other requirements. The computer would then print out these clusters effectively breaking the problem down into independent sub-problems each relatively simple for the designer to understand and solve.

Alexander's work has been heavily criticised, not least by himself, (Alexander 1966) and an excellent review of many of its failings is to be found in Broadbent's book *Design in Architecture* (1973). Some of Alexander's most obvious errors, and those which interest us here, result from a rather mechanistic view of the nature of design problems which is enshrined in much housing legislation which is still in force today. Alexander summarises his attitudes towards design problems as "the problem is defined by a set of requirements called M. The solution to this problem will be a form which successfully satisfies all of these requirements." Implicit in this statement are a number of notions now commonly rejected (Lawson 1979). First, that there exists a set of requirements which can be exhaustively listed at the start of the design process. As we saw in chapter 3, this is not really feasible since all sorts of requirements are quite likely to occur to designer and client alike



5.8 Alexander's method sought to identify clusters of relatively independent requirements which could be considered separately

even well after the synthesis of solutions has started. The second misconception in Alexander's method is that all these listed requirements are of equal value and that the interactions between them are all equally strong. Common sense would suggest that it is quite likely to be much more important to satisfy some requirements than others, and that some pairs of requirements may be closely related while others are more loosely connected. Third, and rather more subtly Alexander fails to appreciate that some requirements and interactions have much more profound implications for the form of the solution than do others.

To illustrate these deficiencies consider two pairs of interacting requirements listed by Chermayeff and Alexander (1963) in their study of community and privacy in housing design. The first interaction is between "efficient parking for owners and visitors; adequate manoeuvre space" and "separation of children and pets from vehicles". The second interaction is between "stops against crawling and climbing insects, vermin, reptiles, birds and mammals" and "filters against smells, viruses, bacteria, dirt. Screens against flying insects, wind-blown dust, litter, soot and garbage." The trouble with Alexander's method is that it is incapable of distinguishing between these interactions in terms of strength, quality or importance, and yet any experienced architect would realise that the two problems have quite different kinds of solution implications. The first is a matter of access and thus poses a spatial planning problem while the second raises an issue about the detailed technical design of the building skin. In a normal design process these two problems would be given emphasis at quite different stages. Thus in this sense the designer selects the aspects of the problem he wishes to consider in order of their likely impact on the solution as a whole. In this case, issues of general layout and organisation would come long before the detailing of doors and windows. Unfortunately the cluster pattern generated by Alexander's method conceals this natural meaning in the problem and forces a strange way of working on the designer.

The rather bald, value-free list of requirements which is the central feature of Alexander's method bears a strikingly close resemblance to the Parker Morris report and subsequent Government Mandatory Minimum Standards for Local Authority Housing. Because the solution implications of these mandatory requirements are so specific it is easy for the designer to find himself designing literally from the kitchen sink, and worktops, outwards. When a requirement is mandatory it has no relative value, just as in Alexander's method, and must be satisfied whatever the sacrifices. Designers working with such legislation must be careful not to act like the student who, when sitting an examination, tends to answer the compulsory question first.

Such design legislation has only recently come under close and critical scrutiny, and designers have begun to report the failings of legislation in practice. In 1973 the Essex County Council produced its now classic Design Guide for Residential Areas, which was an attempt to deal with both qualitative and quantitative aspects of housing design. Visual standards and such concepts as privacy were given as much emphasis as noise levels or efficient traffic circulation. Whilst the objectives of this and the many other design guides which followed were almost universally applauded, many designers have recently expressed concern at the results of such notes for guidance actually being used in practice as legislation. The national building regulations have come under increasing criticism from architects who have shown how they often create undesirable results (Lawson 1975) and proposals are now being put forward to revise the whole system of building control (Savidge 1978).

In 1976 the Department of the Environment published its research report no 6 on the Value of Standards for the External Residential Environment which concluded that many currently accepted standards were either unworkable or even positively objectionable. The report firmly rejected the imposition of requirements for such matters as privacy, view, sunlight or daylight. "The application of standards across the board defeats the aim of appropriately different provision in different situations." This report seems to sound the final death knell for legislation based on the 1960's first generation design methodology. "The qualities of good design are not encapsulated in quantitative standards ... It is right for development controllers to ask that adequate provision be made for, say, privacy or access or children's play or quiet. The imposition of specified quantities as requirements is a different matter, and is not justified by design results."

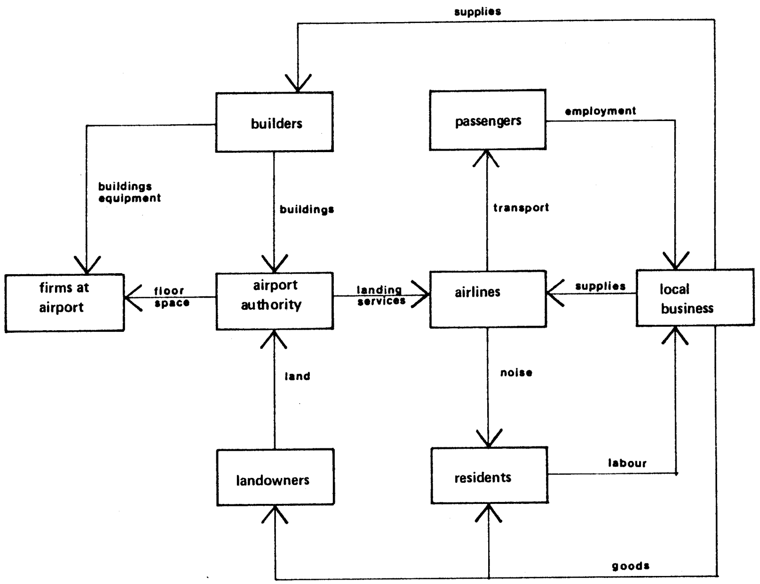
Perhaps it is because design problems are often so intractable and nebulous that the temptation is so great to seek out measurable criteria of satisfactory performance. The difficulty for the designer here is to place value on such criteria and thus balance them against each other and factors which cannot be quantitatively measured. Regrettably numbers seem to confer respectability and importance on what might actually be quite trivial factors. Boje (1971) in his book on open-plan office design provides us with an excellent demonstration of this numerical measuring disease. He calculates that it takes on average

about 7 seconds to open and close an office door. Put this together with some research which shows that in an office building accommodating 100 people in 25 rooms on average each person will change rooms some 11 times in a day and thus, in an open plan office Boje argues, each person would save some 32 door movements or 224 seconds per working day. Using similar logic Boje calculates the increased working efficiency resulting from the optimal arrangements of heating, lighting and telephones. From all this Boje is then able to conclude that a properly designed open-plan office will save some 2,000 minutes per month per employee over a conventional design.

The unthinking designer could easily use such apparently high quality and convincing data to design an office based on such factors as minimising "person door movements". But in fact such figures are quite useless unless the designer also knows just how relatively important it is to save 7 seconds of time. Would that 7 seconds saved actually be used productively? What other, perhaps more critical, social and interpersonal effects result from the lack of doors and walls? So many more questions need answering before the simple single index of "person door movements" can become of value in a design context.

Because in design there are often so many variables which cannot be measured on the same scale value judgements seem inescapable. For example in designing electrical power tools convenience of use has often to be balanced against safety, or portability against robustness. Although it may prove possible to measure designs on crude scales of satisfaction for each of these factors, they remain difficult to relate. Thus a very lightweight lawn mower while being easy to manoeuvre and push might also prove to be noisy and easily damaged. For such an item there is no one right answer since different purchasers are likely to place different values on factors such as manoeuvrability or reliability. The sensible manufacturer of such equipment will produce a whole range of alternative designs each offering different advantages and disadvantages. The problem of relative values however becomes much more critical when design decisions are being taken for large numbers of people who may not have the choice available to the purchasers of new lawn mowers. Examples of such design problems include public sector housing or a new school, the routing of new roads or the siting of factories. Inherently, such projects involve varying degrees of benefit to some and losses to others. A new motorway may well save a long distance motorist's time and relieve congestion in nearby towns while subjecting local residents to noise and pollution.

Attempts have been made to apply cost-benefit analysis techniques to



5.9 A simplified diagram of some of the effects various interested parties have on each other when a new airport is built

these kinds of design problem. Cost-benefit analysis relies upon expressing all factors in terms of their monetary value, thus establishing a common metric. Unfortunately, some factors are rather more easily costed than others. This is perhaps best illustrated by reference to one of the most well known applications of cost-benefit analysis, the Roskill Commission on the siting of the third London airport. After a number of preliminary stages during which some seventy-eight sites were considered the commission narrowed the choice down to four sites at Cublington, Foulness, Nuthampstead and Thurleigh which were then compared using cost-benefit analysis. Even the grossly simplified diagram reproduced here gives some idea of the complex array of effects which the various interested parties could be expected to have on each other as a result of such a project. In fact there are many other much wider effects not shown which include such matters as the distortion of the national transportation network resulting from the provision of new forms of access to the chosen site. For example the opening of an airport at Cublington would have resulted in the closure of the existing Luton airport which would have been too close for air traffic control procedures.

Many of the benefits of the airport in terms of the profits to the various transportation authorities and other companies were reasonably easy to calculate for each site and could be set against the profits lost from the existing use of land. The costs of providing the access transportation to each site and the costs in terms of journey time were also fed into the equation. Losses in terms of reduced amenity however proved more difficult to assess in purely monetary terms. These effects range from otherwise unwanted expenditure resulting from people having to leave their homes, through such factors as the depreciation in value of property in the surrounding area to the noise annoyance caused by the operation of the airport. Such a public use of cost-benefit analysis revealed many of the real dangers involved in basing decisions on the quantification of qualitative factors such as the amenity of an environment. Obviously the success of such a process is contingent upon the assumption that all the costs of amenity loss have been correctly valued. The real difficulty here is that such valuations are unlikely to be arrived at by consensus in a pluralistic society. The costing of noise annoyance or the value of quiet had proved difficult enough for the Roskill Commission, but when considerations of the conservation of wildlife at Foulness were introduced to the argument the whole decision making process began to split at the seams. Cost-benefit analysis was clearly incapable of developing one equation to balance the profits of an airport against the loss of a totally unproductive but irreplacable and, some would say, priceless sanctuary for birdlife. The Roskill report itself recognised the futility of attempting totally objective judgement in comparing the Cublington and Foulness sites. The choice was between the damage to the value of Aylesbury and the loss of a fine Norman church at Stewkley or the ruining of the Essex coastline and probable extinction of the dark-bellied Brent goose.

> As with much else in this inquiry there is no single right answer however much each individual may believe there is. For us to claim to judge absolutely between these views (the importance of conservation of buildings or wildlife) is to claim gifts of wisdom and prophecy which no man can possess. All we can do is respect both points of view. (Roskill Commission Report).

Even the costings of the more ostensibly easily quantifiable factors proved extremely debatable. For example the cost-benefit research team itself revised the assumptions on which total construction costs had been based. This change proved so drastic that Cublington moved from being the most costly to the least costly of the few sites in this respect. As the inquiry proceeded it gradually became apparent that many of the fundamental underlying assumptions necessary for the cost-benefit analysis could similarly be challenged. Basing the whole decision making process on arguments developed from such questionable assumptions proved too uncomfortable for at least one member of the commission. In his minority report Professor Buchanan described how "I became more and more anxious lest I be trapped in a process which I did not fully understand and ultimately led without choice to a conclusion which I would know in my heart of hearts I did not agree with."

In the final analysis it seems unreasonable for designers to expect to find a process which will protect them from the painful and difficult business of exercising subjective judgement in situations where both quantitative and qualitative factors must be taken into account. The attempt to reduce all factors to a common quantitative measure such as monetary value frequently serves only to shift the problem to one of valuation. The Roskill Commission on the siting of the third London airport provided one further lesson of importance here. Designers and those who make design-like decisions which profoundly affect the lives of many people can no longer expect their value judgements to be made in private. Such large-scale design processes must clearly invite the participation of all those who will be substantially affected. How that participation should be organised is another matter beyond the scope of this particular book.

6 A model of design problems

As an artist I did not set out to make the public understand but to find problems for myself of space and form, and to explore them.

Henry Moore (on his 80th birthday)

Open-ended problems are both easy and difficult to solve. They are easy in that even untutored persons can arrive at a solution, but difficult when the solver sets forth a comprehensive list of criteria to which his solution must conform and, whether implicitly or explicitly he attends to this list in his solution.

Robert Wehrli, *Open-ended Problem Solving in Design*

The generators of design problems

It is sometimes difficult to separate design from art. The products of design are frequently seen by the public as artistic, even sometimes actually as "works of art", and designers themselves are indeed also often artists. Even the drawings generated by designers to illustrate their schemes can sometimes easily be confused with works of art. Whether or not an object can rightly be described as a "work of art" is a matter which lies beyond the scope of this book. What is of importance here is not the product but the process. The creative process which may give rise to a work of art undoubtedly shares much in common with the design process, and many of the same talents may be needed for both. What is usually different however is the nature of the source of the problem. Perhaps it is only a difference of degree, but nevertheless the difference is real enough and of fundamental significance in determining the nature of the processes of art and design. While the artist may sometimes be commissioned to produce work for a particular place or occasion he is more often entirely his own master. The artist deals with issues and solves problems which seem important to him. The artist may respond to his own work and is free to shift his attention and change the problem to one which fascinates him more. Such artistic problems are rarely clearly articulated by the artist outside his work. It is usually critics and historians who retrospectively interpret and identify