Decision-Making: A Laboratory Based Case Study in Conceptual Design

by
Maurice Girod

A Doctoral Thesis

Submitted in partial fulfilment of the requirements
for the award of
Doctor of Philosophy of Loughborough University

August 2001

Acknowledgements

It is important to me to acknowledge those who have made this project possible and those who have made it enjoyable.

First of all I would like to thank the Engineering and Physical Sciences Research Council (EPSRC) as well as Loughborough University, Faculty of Engineering, for providing the necessary finances. Thanks are also due to all those individuals who participated in the research.

I wish to express my gratitude and indebtedness to my supervisors, Prof. Ian Wright and Prof. Neil Burns, for their guidance and support and for their persistent willingness to read, re-read and re-re-read my reports, thesis chapters and papers. In this respect I would also like to thank my invaluable colleague Amanda Elliott especially for all her efforts in helping to make my writings readable, but also for her constructive criticism on the content.

There were two particular friends who helped me throughout the project in various ways: I will never forget the enlightening and very enjoyable discussions about life, the universe and everything with John McGuire. Neither will I forget the sheer endless help supplied by John Jones when it came to get my computer running again and when my poor old car demanded repeated attention.

I am also grateful to my parents, Gisela and Wolfgang, for sending survival parcels and for never giving up hope that I would eventually complete the project.

Finally, very special thanks go to Michaela. Without her constant encouragement and her efforts to make me think 'big picture' I would probably still be lost in some hidden corner of the project. I do not want to forget mentioning her delicious nutritional inputs either.

Abstract

The engineering design process may be seen as a series of interrelated operations that are driven by decisions: each operation is carried out as the consequence of an associated decision. Hence, an effective design process relies heavily upon effective decision-making. As a consequence, supporting decision-making may be a significant means for achieving design process improvements. This thesis concentrates on how to support *selection-type* decision-making in *conceptual* engineering design.

A literature review showed that there is a variety of available methods that may all be used to support selection-type decision-making in conceptual engineering design. However, the review indicated that, although there is an apparent need, the methods are not used widely and that very little is known about support requirements.

The main research effort of this study aimed at identifying requirements for decision-making support. To do so, new insight about how decision-making processes actually take place was gathered. Workshops involving different groups being engaged in conceptual design decision-making processes were observed, recorded and transcribed. Subsequently, the transcripts were analysed using content analysis and pattern coding. This resulted in (i) the identification of a set of decision-making related activity categories, (ii) the quantification of each category's time consumption, (iii) the identification of different formats of particular information and (iv) the generation of a detailed decision-making process model.

The study's results suggested various implications. These are expressed as 'observations' on specific process characteristics and finally as requirements for decision-making support. Overall, it is suggested that effective decision-making support for engineering design involves implementing a flexible, computerised information management system. The specific requirements identified by this study may be applied for directing the development of such a system.

Keywords: engineering design, conceptual design, decision-making, methods, activity categories, process model, empirical research.

List of Publications Relevant to this Thesis

Girod, M., Elliott, A.C. and Wright, I.C., 2000a, 'Decision-making and design concept selection', *Proceedings of the Engineering Design Conference, EDC2000*, Brunel, UK, pp.659-666.

Girod, M., Elliott, A.C., Wright, I.C. and Burns, N.D., 2000b, 'A Descriptive Model of Collaborative Concept Selection Processes in Engineering Design', *Proceedings of the International Conference on Concurrent Engineering, CE'00*, Lyon, France, pp.494-503.

Girod, M., Elliott, A.C., Wright, I.C. and Burns, N.D., 2000c, 'Activities in Collaborative Concept Selection Processes for Engineering Design', *Proceedings of the ASME-DETC'00 Conference on Design Theory and Methodology*, Baltimore, USA, paper: DETC2000/DTM-14548.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The level of competition in the manufacturing industries can be described as being fierce. Companies seek to gain advantage not only by producing effective products, but also by introducing them to the market earlier than their competitors. They try to improve their product design processes in order to drastically decrease both the product lead times and costs under increasing quality standards - their challenge is to produce faster, cheaper and better (Nichols, 1992).

There are a number of methods and tools available that may support engineering designers. However, most of them focus on the design artefact rather than on the process of designing it; thus, there is a need for more research directed generally towards the improvement of design processes and more specifically directed towards the development of methods and tools (Wallace & Burgess, 1995).

The engineering design process may be defined as a sequence of decisions (Marples, 1960) or else as a series of interrelated operations that are driven by decisions (Midland, 1997). This means that each operation is carried out as the consequence of an associated decision. Hence, an effective and efficient design process relies heavily upon effective and efficient decision-making. As a consequence, supporting decision-making can be a significant means for achieving design process improvements.

In design, a decision can be defined as the selection of one alternative from a number of alternatives (Colton & Pun, 1994). Such a selection is based on the alternatives' effectiveness with respect to evaluation criteria. Virtually all engineering design decisions are based on multiple criteria (Roozenburg & Eekels, 1995). The field of research concerned with supporting multiple-criteria decision-making is extremely versified. It incorporates various aspects such as: context-related as well as structure-related support; psychological support, technological support, and methodological support. Focusing on the last aspect there are: methods for the establishment of criteria importance weightings and methods for determining the best solution; optimisation methods and selection methods; methods that do not model uncertain information and methods that do model uncertain information. Explanations of these aspects will be delivered in section 1.3. The particular focus of this thesis is structure-related, methodological support for selection problem situations during the *conceptual* engineering design phase. In short: general methods that support the selection of design concepts. This focus will be justified in section 1.3.

The literature review in chapter 2 shows that a number of methods have been developed in academia that may be of benefit for supporting the resolution of selection problems in conceptual design. However, the effectiveness of these approaches has not been sufficiently researched and it is claimed by practitioners that they are not used widely in industrial practice (Schlüter, 1999b). It is difficult to assess whether or not the characteristics of these methods actually match the requirements of effective decision-making support for conceptual engineering design. This is due to an apparent lack of research to identify such requirements. It is therefore worthwhile to step back and dedicate more work towards identifying practitioners' needs, i.e. the requirements for decision-making support. This sets a practical basis upon which available methods may be critically assessed and it helps to direct efforts towards the development of new decision-making methods or tools.

This thesis reports on a project that addresses the overall question: how to support selection-type decision-making in conceptual engineering design. The aims are to develop an understanding of the research area and to identify requirements for decision-making support. An understanding of the research area was gained through discussing the characteristics of available support methods. Requirements for such methods were then identified by empirically investigating decision-making processes. This investigation comprised two main stages: first, establishing descriptive process models; and second, analysing the models.

1.2 Research problem, questions and objectives

The general research problem addressed by this thesis is: How to support selection-type decision-making in conceptual engineering design?

As a contribution towards resolving the general research problem this thesis delivers answers for the following research questions:

1. What are the specific characteristics of existing decision-making support methods?
   Finding an answer to this question involved the following objectives:
   1.1 Identify typical decision-making support methods;
   1.2 Discuss the methods with the aim of understanding their characteristics.
2. What are the requirements for decision-making support?
   Finding an answer to this question involved the following objectives:
   2.1 Develop a detailed understanding of actual decision-making processes;
   2.2 Identify support requirements through the developed understanding.
3. Do the identified methods' characteristics match the identified requirements?
   Finding an answer to this question involved the following objective:
   3.1 Compare the requirements with the methods' characteristics.

For question 1 I have reviewed the literature on methods for decision-making support. For question 2, which took the main research effort, I have conducted an empirical investigation on decision-making processes in conceptual design. For question 3 I have reflected on the results of my empirical investigation and considered them in relation to the answers that I have found for question 1. The research questions and objectives are discussed in more detail in chapters 2 and 3.

I conclude that the available methods, at least those identified and discussed in chapter 2, do not fully satisfy the support requirements for the decision-making situation that I investigated. This will be shown by inductive argumentation in chapter 5. Therefore, I suggest that more research into the development of new support approaches, methods and tools specifically for conceptual design decision-making is required. As guidance for such research and development I propose a set of support requirements.

The study described by this thesis contributes to the body of knowledge in the specific research area, i.e. 'decision-making support in conceptual engineering design', mainly by providing new insight into how decision-making processes actually take place. In particular this is the empirical establishment of a new decision-making process model at a very detailed level and the proposition of a new set of requirements for decision-making support based on empirical research.

1.3 Justification for the research

The study *generally* aimed at improving engineering design processes. Specifically I chose to focus on decision-making support, on conceptual design and on selection-type decision-making. I also chose to largely base the project on empirical research. In the following I will deliver justifications for these choices.

1.3.1 Focus on decision-making support

It is a major challenge for manufacturing companies not only to produce better products, but also to produce them cheaper and faster than their competition (Nichols, 1992). To do so, the design processes need to be effective and efficient. These processes are seen as sequences of decisions, as argued in section 1.1. Therefore, effective and efficient design processes rely heavily upon effective and efficient decision-making. In the quest for improvements books are sold, articles published and conferences organised that promise to help engineering designers to make decisions more effectively and more efficiently (Herling, 1997). This suggests that the engineering design community is indeed interested in facilitating decision-making.

A number of methods have been developed in academia that may be of benefit for supporting decision-making. Yet, they are not used widely in industrial practice (Schlüter, 1999b). On the other hand it has been realised that decision-making in engineering design causes problems in many practical cases (Weiss & Hari, 1997) and that much product development time is wasted by making poor decisions

(Ullman, 2000). This indicates that practitioners could indeed benefit from more research on how to support their decision-making processes.

1.3.2 Focus on conceptual design

There are a number of different design process models in the literature. Basically, these models divide the design process into distinct phases (Roozenburg & Eekels, 1995). Predominant examples of such phase models have been developed by French (1985), Pahl and Beitz (1988) and Pugh (1990). Appendix A shows these models in graphical form. In Pahl and Beitz's (1988) model there are four phases called:

1. Clarification of the task;
2. Conceptual design;
3. Embodiment design;
4. Detail design.

The first phase 'clarification of the task' involves analysing a given problem and generating a product design specification. Such a specification may then direct the actual design work in the following phases of the design process. The second phase 'conceptual design' involves generating and evaluating broad solutions. Such solutions are called 'concepts' (Pahl & Beitz, 1988) or 'schemes' (French, 1985). These broad solutions provide the point of departure for embodiment and detail design. The third phase 'embodiment design' involves elaborating the selected concept into a definite design. The definite design determines the layout of assemblies, components, shapes, rough dimensions and materials. The fourth phase 'detail design' involves fully specifying all the details roughly determined in the previous phase and producing the product documents.

In accordance with the phases of the actual design work, i.e. phase 2, 3 and 4, as discussed above, Starkey (1992) categorised the overall importance of design decisions. He categorised them on three different levels; fundamental decisions, intermediate decisions and minor decisions, and explains their meaning as follows:

Fundamental decisions: In any design process there will be a small number of decisions which have tremendous consequences for the whole design project and can determine the success or failure of the product. These decisions are made at the early stages of the design process when the designer has little concrete knowledge about the way ahead. However, the basic concepts of the product have to be chosen. This type of decision is called fundamental decision. Changing a fundamental decision often means beginning again from scratch.

Intermediate decisions: Intermediate decisions stem from fundamental decisions. They may be seen as an extension or a supplement to the fundamentals. The importance of intermediate decisions can differ. Some of them might be nearly as important as fundamentals whereas others may only have little more significance than minor decisions. However, a change of an intermediate decision still means in most cases a considerable amount of re-design.

Minor decisions: Minor decisions stem from intermediate decisions or other minor decisions and occur in a vast number. They are of relatively little importance and can be found mainly in the detail design stage.

The gravity of the fundamental decisions, made during the conceptual phase, is also reflected by the realisation that a large part of a product's cost is disposed already in this phase, although relatively little costs have occurred so far (Andreasen & Olesen, 1990). There are quantified versions of this statement, which may be questioned (Barton et al., 2000), but the considerable importance of early design decisions is still acknowledged, specifically due to their product development downstream impact. As a conclusion, I focused on decision-making in the conceptual design phase.

1.3.3 Focus on selection-type decisions

Decision-making support is a confusingly wide field of research. A large number of publications are available that approach the subject from various aspects. Figure 1.1 shows a taxonomy that addresses some dominant aspects and clarifies how they relate to one another. I developed this taxonomy, which reuses earlier published frameworks (Sen & Yang, 1995; Zimmermann, 1996), to gain an overview of the general field 'decision-making support' and to focus the project. Retrospectively it allows for seeing the focus of this project in a larger frame of reference.

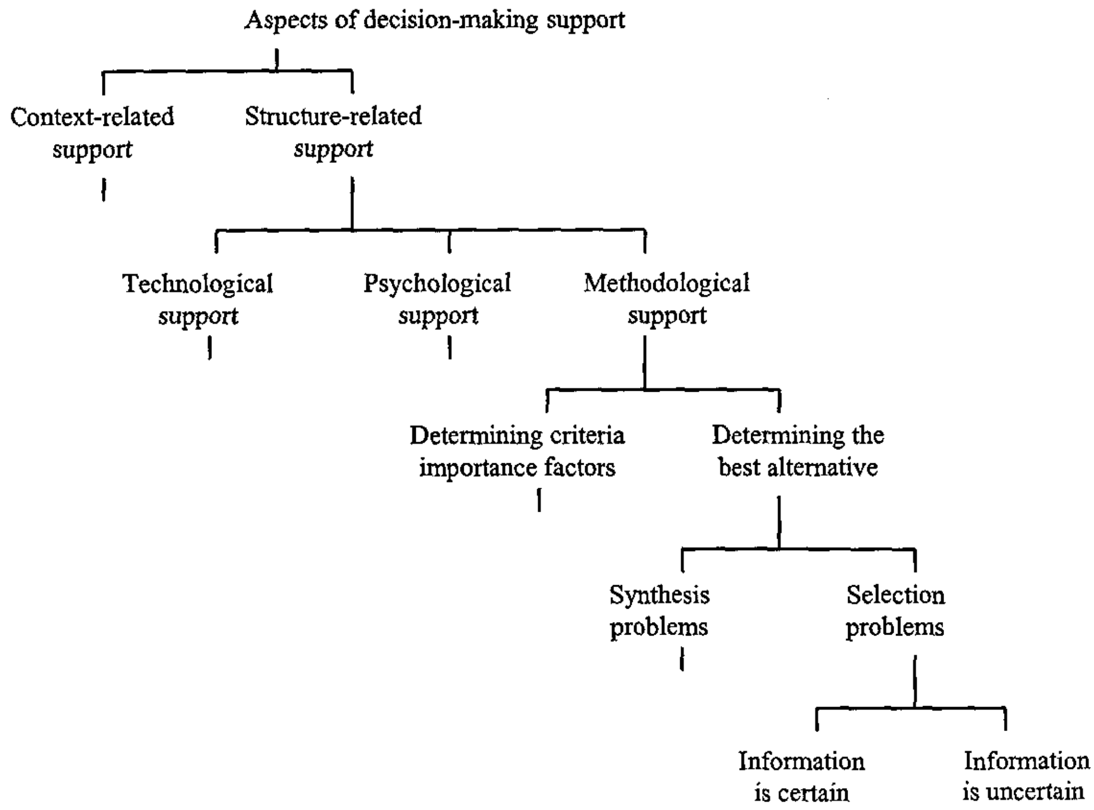


Figure 1.1: Aspects of decision-making support.

A main aspect for the differentiation of decision-making support approaches is their context. *Context-related support* approaches, such as expert systems, are limited to decision support with respect to a particular area of application. In contrast, *structure-related support* approaches are context independent. Because the former have much smaller domains of application than the latter (Zimmermann, 1996), I concentrated on the latter, i.e. *structure-related support.*

Structure-related support involves *technological support, psychological support,* or *methodological support. Technological support* can be almost any technology that contributes towards making decisions in the widest sense. Examples are communication support tools and electronic meeting systems, which are often referred to as group decision support systems (Choi et al., 1994). *Psychological support* is concerned with human behaviour in specific situations, in particular with respect to human interaction. *Methodological support* concentrates on structuring decision situations. I was interested in the very basic decision-making support principles, which may or may not be supplemented by technological or psychological aspects. I found that these basic principles fall under *methodological support* – directing my further investigations towards this aspect.

Methodological support often addresses the evaluation of decision alternatives with respect to decision criteria that are of unequal importance. Determining the importance of decision criteria, especially by assigning factors, is not always easy. There are approaches that particularly support the task of *determining criteria importance factors*. Other *methodological support* approaches concentrate on *determining the best alternative* from the so-called 'alternative space'. This involves aggregating information that refers to the alternatives' performances with respect to the individual decision criteria. Such aggregations result in statements on the alternatives' overall performance. Some, but not all approaches apply aggregations that consider importance factors for decision criteria. Therefore, I see the task of *determining criteria importance factors* as a more specific aspect of decision-making than generally *determining the best alternative.* Seeking generality, I focused on the latter aspect.

Approaches that aim at *determining the best alternative* can be classified according to the broad type of decision problem to which they refer. Two occurring problem classes are *synthesis problems* and *selection problems* (Sen & Yang, 1995): synthesis involves optimising an alternative by adjusting a number of parameter values, which is called multi-objective decision-making (MODM); selection involves choosing one discrete alternative from a number of available discrete alternatives, which is called multi-attribute decision-making (MADM). Even though the literature classifies optimisations as multi-objective *decision-making* it is arguable whether optimisations actually refer to *decision* problems at all. Purists may take the view that optimisations are pure mathematical exercises with little or no human judgement being required and as such do not really involve making a *decision* (Jebb & Woolliams, 2000). A number of researchers in the field of engineering design have generally defined decisions as the *selection* of one alternative from a number of alternatives (Pahl & Beitz, 1988; Colton & Pun, 1994), which is clearly distinct from optimisations as for solving synthesis problems. Also, it has been reported that in particular the *selection* of concepts in the conceptual design phase, upon which this research focuses, causes difficulties in many practical cases (Weiss & Hari, 1997; Nidamarthi, 1999). As a conclusion I focused this project on support for *selection problems*.

A number of approaches for *selection problems* assume that the available *information is certain*. Yet, in engineering design this assumption is not always valid, i.e. there are situations in which uncertainties need to be taken into account (Sen & Yang, 1995). There are specific approaches applicable for cases in which *information is uncertain*. I have taken this aspect into consideration.

1.3.4 Empirical research

The literature review revealed that various methods have been developed in academia that may be applied to support selection-type decision-making in conceptual design. However, the literature review also revealed that there is an apparent lack of empirical research aiming at (i) understanding how selection-type decision-making processes actually take place in conceptual design and (ii) finding the requirements for support. This lack of research makes it difficult to assess the effectiveness of available methods for conceptual design applications and to direct further work towards the development of new support approaches.

Three relevant previous empirical projects could be identified in the literature (see chapter 2). One of them focused on studying selection-type decision-making processes, but did not aim at identifying support requirements. Instead, it was aimed at studying the influence of various parameters onto evaluation processes. No detailed decision-making process models were generated. The other two projects did aim at identifying decision-making support requirements, but focused on studying 'overall' conceptual design processes. No detailed decision-making process models were generated. In contrast, the project described by this thesis, focuses on studying selection-type decision-making processes, rather than 'overall' design processes, with the aim to identify support requirements through establishing a detailed decision-making process model.

1.4 Methodology

The research project described by this thesis was of an exploratory nature: it explored decision-making processes with the aim of increasing the understanding about such processes. I set the project in the constitutive research paradigm. Therefore, the gathered data is mainly qualitative and it was not intended for this project to make generalisations. Previous research (e.g. Ullman et al., 1988; Dwarakanath, 1996; Nidamarthi, 1999) and methodological considerations by Cannon and Leifer (1999) suggest that exploratory research set in the constitutive research paradigm is an accepted practice for engineering design investigations.

To find out about the particular characteristics of available decision-making methods, i.e. to address research question 1, I conducted a literature review. The databases that I mainly used for this review were BIDS (http://www.bids.ac.uk) and OCLC (http://www.oclc.org).

To find out about the requirements of decision-making processes, i.e. to address research question 2, I conducted empirical research. The use of questionnaires or formal interviews, as often applied for such type research, did not seem ideal data collection methods for my endeavour. This is because questionnaires and also formal interviews have potential limitations on data due to the restrictions they place on the depth of questioning (Mutchnik & Berg, 1996). It would have been possible to conduct informal interviews as they allow for more depth of questioning. Yet, the preferred method for data collection was observation through ethnography, which takes the researcher directly to the process under investigation (Berg, 1989) rather than the other way round. By this means a comprehensive description of details regarding the investigated process may be gained (Mutchnik & Berg, 1996). However, initial attempts to implement the ethnographic approach failed due to access problems. The reason was that organisations considered concept phase decision-making meetings to be too commercially confidential to allow observation and recording.

Due to the problems with the ethnography I applied an alternative approach, still using observation. I organised decision-making workshops conducted under laboratory conditions within the university. Apart from practicality, this had the advantage of reducing the study's complexity and increasing the level of control, i.e. to keep certain variables constant across a number of workshops. These variables referred to the environment in which the decisions were taken and the context, i.e. the available alternatives and criteria. University based observational studies were also used for design research in previous projects (Dwarakanath, 1996; Nidamarthi, 1999).

Altogether I organised five workshops, two of them being pilots to test set-up and equipment. During these workshops I observed groups whilst being engaged in collaborative decision-making processes. They were all given the same task, to evaluate a number of conceptual design solutions and to select, as a group, the most effective one. The design solutions were not developed during the workshops, but had been produced in a number of preceding brainstorming sessions, which had involved the workshop participants. These design solutions were different conceptual alternatives for a portable device that enables one operator to place full CO_2 cylinders into the cellar of a Public House (pub) and raise empty cylinders. A diagrammatic layout of the pub and cellar is given in figure 1.2. I recorded the workshops and fully transcribed them for later analysis.

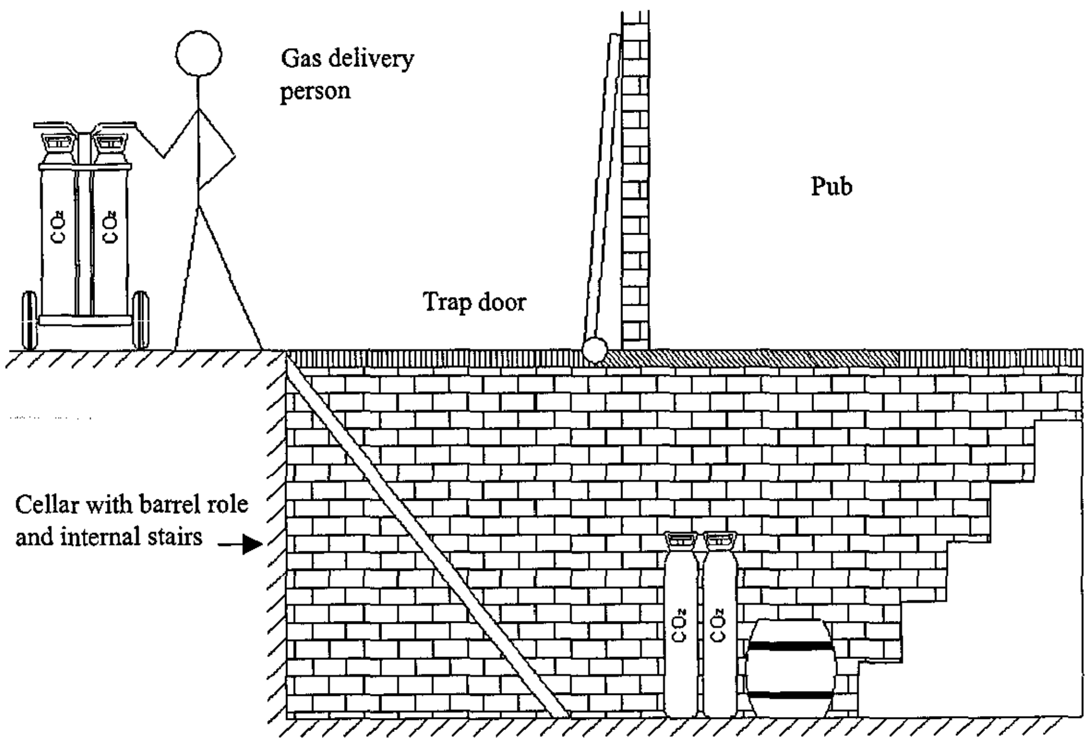


Figure 1.2: Diagrammatic general layout of pub and cellar.

The chosen design issue, as outlined above, was a 'real' problem of a major British industrial company. The suitability of this particular design issue and its alternative solutions for my research will be discussed in section 3.5.1.

When analysing my data, i.e. the transcripts, I first established a set of decision-making related activities that were carried out by the observed groups. I then used the identified activities to generate detailed decision-making process models. Measuring the activities' time consumptions as well as generating the process models allowed me to make a number of observations and suggested requirements for support.

To find out whether the support methods' characteristics match the suggested requirements, i.e. to address research question 3, I reflected on my results in a discussion. Chapter 3 will introduce the research methodology in greater detail.

1.5 Delimitations of this study

This study has a number of deliberate limits. These limits refer in particular to the research context and the generalisability of the results.

The context-related limits refer to the specific subject area, which is engineering design and in particular conceptual design and to the type of decision-making problem which is selection-type decision-making. Any identified characteristics of decision-making methods and any claims that the available decision-making support methods do not fully satisfy practical requirements are restricted to those methods that could be identified by the literature review as in chapter 2.

The limits related to the generalisability of the results are due to the explorative nature of the study and in particular due to the small sample size of observed subjects/groups. This study's results may therefore be seen as hypotheses that may be tested and generalised by subsequent quantitative research. The project's limits have been justified in section 1.3 and 1.4.

1.6 Thesis structure

This thesis is divided into five main chapters plus one chapter summarising the conclusions. The structure is shown in figure 1.3. Chapter 1 introduces the background, states the research problem and outlines the research questions and objectives as well as the research methodology. It also delivers justifications for the research problem and the methodology. Further, the delimitations of the study are shown.

Chapter 2 reports on the literature review. Typical methods for decision-making support will be discussed with the aim to identify their particular characteristics. It will be shown that the field has major contributions from Operational Research, Artificial Intelligence research and engineering design research. The identified previous projects that have empirically studied decision-making processes in conceptual design will also be discussed in this chapter. This latter discussion leads to the identification of an important variable to be considered in empirical studies on decision-making: the decision-makers' level of professional experience.

Chapter 3 describes in detail the methodology applied for data collection and data analysis of the empirical investigation. Not only the theoretical aspects will be

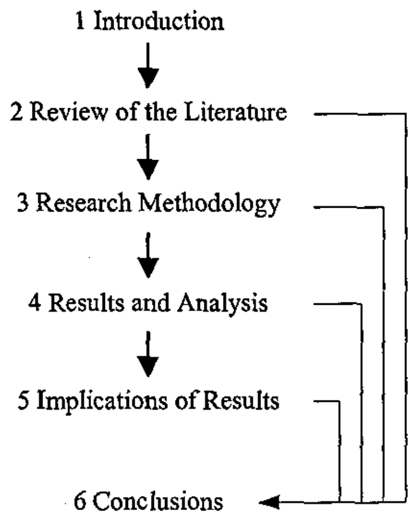


Figure 1.3: Thesis structure.

discussed, but also there will be room for deliberating some very practical issues to be considered for conducting such investigations.

Chapter 4 reports on the results from the data analysis. These are a set of decision-making related activities as well as their time consumptions and a detailed decision-making process model.

Chapter 5 discusses the implications of the results, which are expressed as observations and finally as a set of requirements for decision-making support. These requirements are then compared with the characteristics of available decision-making support methods, as identified in chapter 2.

Chapter 6 summarises the main conclusions from my study. This chapter also points out the study's contributions as well as its limitations and also suggests some directions for short term as well as long term further work.

1.7 Summary

This chapter laid the foundations for the thesis. It introduced the research problem and main research questions. Then the contribution to the body of knowledge was indicated and the research was justified. The methodology was briefly described and justified, the delimitation was given and the thesis was outlined. On these foundations, the thesis can proceed with a detailed description of the research.

Chapter 2

Review of the Literature

2.1 Introduction

To facilitate decision-making three things are necessary: options, expectations, and values (Hazelrigg, 1997). Options may be alternative design solutions. Their number is mostly unlimited (Ullman & D'Ambrosio, 1995). Expectations are predictions about the performance of the alternatives. In conceptual design the alternatives are not actually existent at the time when a decision on them has to be made. Therefore, there is the inevitable problem of accurately forecasting the future (Efstathiou, 1984). This may imply that information on the alternatives' performance is uncertain (Sen & Yang, 1995; Hazelrigg, 1997; Carnahan et al., 1994). Values refer to what the decision-maker wants and enable the preference ranking of the expectations of different alternatives for a particular decision (Hazelrigg, 1997). The decision-maker's values determine the goals which are expressed through criteria and associated weights. The criteria refer to a usually large number of attributes, which the alternatives are desired to possess (Efstathiou, 1984; Sen & Yang, 1995).

The alternatives are evaluated with respect to their criteria-related attributes or, in other words: the level of satisfaction that one gains out of a particular grade of the criteria's fulfilment. Many of the desired attributes are conflicting (Efstathiou, 1984) so that a trade-off between them has to be found (Yang & Sen, 1997). This trade-off can be modelled by a combination function, such that some attributes of an alternative may contribute more heavily to the combined, overall, evaluation result than others (Antonsson & Otto, 1995). This means, some attributes will have a stronger influence on the final decision than other attributes. Combination functions are sometimes referred to as aggregation functions, objective functions, or as metrics.

There are two distinct ways of dealing with decision problems. One way is making a more or less *intuitive decision* by applying rules of thumb (heuristics). These rules simplify complex decision problems down to a level that meets the capacity of a human being. The capacity of a human being is limited in particular with respect to simultaneously processing information. This is relevant to decision-making as decision-makers are often required to consider information regarding various decision criteria simultaneously.

The other way of dealing with a decision problem is using a formal method. These methods help the decision-maker by structuring the problem in such a way that the full complexity can be dealt with.

This chapter is to a large extent a discussion on various available methods that may be applied to support selection-type decision-making in conceptual design. The general aim of this discussion is to identify available support methods and to develop an understanding of the research area. An exhaustive overview of all available methods will not be given, but an attempt will be made to generate insight into the 'manner of thinking and reasoning' i.e. the assumptions and the underlying principles upon which the 'mainstream' methods are based. I call these assumptions and principles the methods' characteristics. Thus, this chapter addresses the project's objectives 1.1 and 1.2, which relate to research question 1 as stated in section 1.2 of the thesis.

The literature review revealed that methodological input for supporting decision-making problems in conceptual design has been contributed by at least three research disciplines: Operational Research, Artificial Intelligence and engineering design research. Therefore, I see 'design selection support' as an intersection of these three disciplines. This is illustrated by figure 2.1.

Operational Research contributed many of the fundamental methods and theories about multi-criteria decision-making (Tebay et al., 1984). Artificial Intelligence researchers have developed a number of methods for the representation of uncertain information, which has traditionally been modelled by using probability theory. Engineering design research has lead to the development of methods specifically for

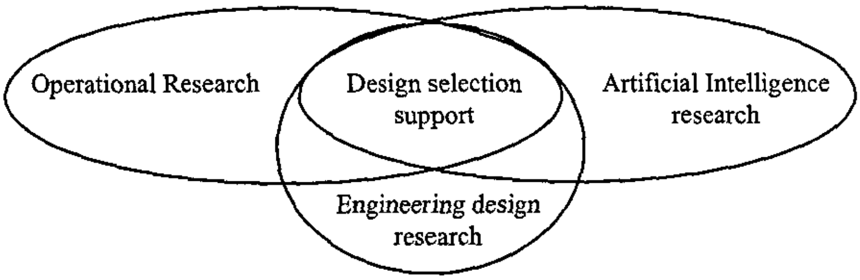


Figure 2.1: The disciplines involved in design selection support.

design selection situations. These methods reuse principles from Operational Research and Artificial Intelligence research.

In the following sections I will first introduce and discuss methods for multi-criteria decision-making including related design selection methods/applications. Then I will introduce and discuss methods for modelling uncertainty including related design selection methods/applications. Finally, I will deliver the conclusions of the literature review. These are a perceived research gap, a research aim, and a number of further, guiding research questions derived from the discussions on the identified methods' characteristics.

2.2 Methods for multi-criteria decision-making

Major multi-criteria decision-making methods in Operational Research are Multi-attribute Value Theory (MAVT), the Analytic Hierarchy Process (AHP) and the outranking methods Electre and Promethee. Major methods specifically developed for engineering design are Pugh's method of controlled convergence and decision-making with QFD. The term 'major' is here related to the frequency of these methods being mentioned in the literature.

2.2.1 Multi-attribute value theory - how it works

MAVT (Keeney & Raiffa, 1976) solves complex multiple-criteria decision problems by disaggregating them into a set of sub-problems, solving each one, and then reassembling them to obtain a solution to the larger decision problem (Weiss & Hari, 1997). The sub-problems are single criterion evaluations, rating the alternatives' performances with respect to individual criteria. The 'overall worth' (Weiss & Hari, 1997) of the alternatives is then found by aggregating the ratings of the several individual performances (Thurston, 1990). MAVT is used for making decisions in an environment of certainty (Simpson, 1996).

For rating the alternatives' performances each of the individual decision criteria must be associated with a function representing the preferences of the decision-maker. This function shows the 'effectiveness' of an alternative's performance with respect to the criterion in question. This means that the alternatives are evaluated on an independent scale, the 'effectiveness function' (Roozenburg & Eekels, 1995), which results in absolute evaluations, ratings. These effectiveness functions reflect the decision-maker's preferences over a range of attribute levels, and incorporate non-linearity of preferences (Thurston, 1990).

Two examples of individual effectiveness functions showing an imagined decision-maker's preferences with respect to the criteria 'maximal speed' and 'fuel economy' of cars are shown in figures 2.2 and 2.3, respectively. If, with reference to the example in figure 2.2, a particular alternative's maximal speed was predicted to be 90 miles per hour, this alternative's effectiveness with respect to max. speed would be 0.7.

In a multi-criteria decision problem each alternative may be evaluated with respect to many criteria. Therefore, every alternative can be associated with a number of effectiveness functions in a multi-dimensional space. That is, each criterion contributes one dimension towards this space. An example of a simple two-dimensional space is given in figure 2.4, which depicts the effectiveness of alternatives A and B, as in figures 2.2 and 2.3.

To determine the overall worth of an alternative, with respect to all criteria, it is necessary to map the multi-dimensional space onto a one-dimensional space. This mapping is achieved by an aggregation function, which performs a trade-off between the various criteria by means of weights. These weights represent the decision-maker's perception of the relative importance of each criterion. The aggregation function usually applied in MAVT is a weighted sum (Antonsson & Otto, 1995), which adds up the products of weight and the alternatives' performance effectiveness for all criteria.

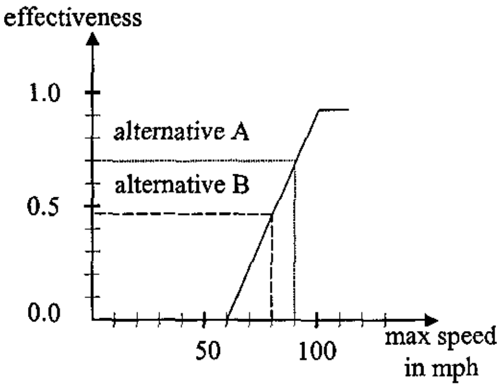


Figure 2.2: Effectiveness of maximal speed.

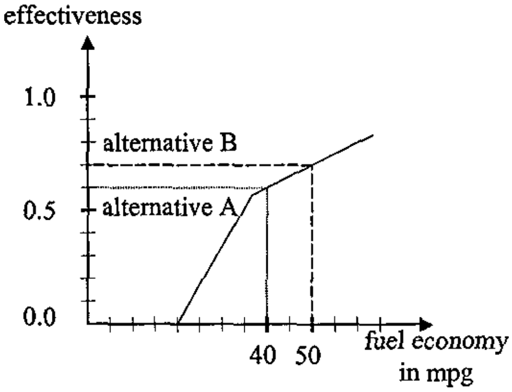


Figure 2.3: Effectiveness of fuel economy.

Rational decisions follow the rule that the overall preferred alternative is the one whose overall worth, as calculated by the aggregation function, is the highest (Hazelrigg, 1998). For the examples in figures 2.2 and 2.3, assuming that maximal speed has a weight of 0.4 and fuel economy has a weight of 0.6, the weighted sums are calculated as follows:

$U(A) = 0.4 \cdot 0.7 + 0.6 \cdot 0.6 = 0.64$

$U(B) = 0.4 \cdot 0.48 + 0.6 \cdot 0.7 = 0.61$

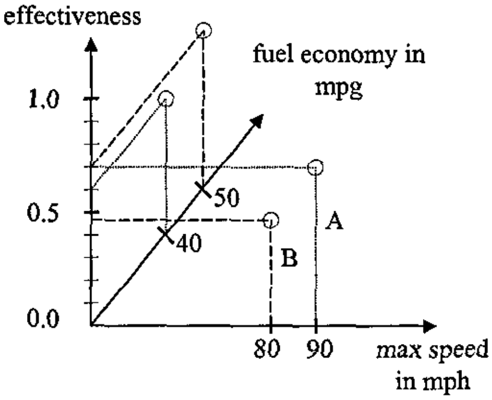


Figure 2.4: Two-dimensional effectiveness.

The result of mapping the two-dimensional space in figure 2.4 onto a one-dimensional space by the weighted sums is shown in figure 2.5.

Application in engineering design

MAVT as introduced above can be applied to engineering design selection problems. This has been shown by Thurston (1990). Her application was the selection of a material system design concept for an automotive frame and body skin.

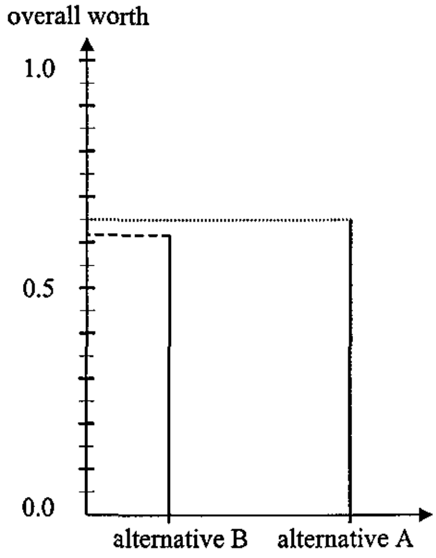


Figure 2.5: Overall worth.

Frequently mentioned in the literature on engineering design are simplified versions of MAVT using directly assigned performance scores (e.g. Roozenburg & Eekels, 1995; Cross, 1994; Pahl & Beitz, 1988). The procedure is to first weight the criteria by assigning a number reflecting their importance; e.g. 10 indicating utmost importance and 1 indicating little importance. Then, the alternatives' performance scores for each criterion are directly determined upon an interval scale, e.g. from 0 indicating no criteria satisfaction to 100 indicating perfect criteria satisfaction. Finally, the overall worth is calculated by applying the weighted sum aggregation function, as discussed above. A very simple, abstract example of such a direct scoring approach is given in table 2.1. This example was taken from Hartvig (1998) who has partly based her prototype decision-support system for conceptual design upon this method.

	Weights	Solution 1	Solution 2	Solution 3
Criterion 1	10	80 · 10=800	60 · 10=600	10 · 10=100
Criterion 2	5	25 · 5=125	30 · 5=150	70 · 5=350
Criterion 3	2	60 · 2=120	70 · 2=140	90 · 2=180
Criterion 4	8	20 · 8=160	50 · 8=400	60 · 8=480
Totals:		1205	1290	1110

Table 2.1: Direct scoring table.

2.2.2 Multi-attribute value theory - characteristics

Underlying assumptions

MAVT has been devised with a disaggregate view of the decision-makers in mind Simpson (1996). This means, they are assumed to be able to examine their own feelings and to be able to communicate their personal preferences and values in a format similar to the inputs required by the method. Due to the application of mathematically rigorous functions MAVT gains a strict axiomatic foundation, which in turn imposes specific conditions for preference inputs (Hazelrigg, 1998). These are:

Transitivity: if a decision-maker prefers A over B and B over C, he/she must also prefer A over C. Hazelrigg (1998) argues that an individual decision-maker may only be considered rational under the condition that his/her preferences are transitive.

Preference independence: each attribute may be evaluated, regardless of the states of the other attributes (Sen & Yang, 1995) and the weight of each criterion must remain constant, whatever the level of fulfilment of the other criteria (Efstathiou, 1984).

Comparability: all alternatives must actually be comparable across all criteria (Simpson, 1996).

It has been argued that at least some of the assumptions underlying MAVT are very strict and cannot actually be conformed to in many practical cases. Some researchers believe that, in particular, the required preference independence cannot always be assumed in practical decision-making situations (Wang, 1994; Sen & Yang, 1995; Simpson, 1996).

Further, the method implicitly assumes that a decision-maker already has a consistent and well-defined system of preferences at the outset of the decision-making process. Vetschera (1994) questions the practicality of this assumption. Wang (1994) believes decomposing a decision-maker's preference structure into mutually independent elements may be extremely difficult especially in complex situations. He thinks that, as an intelligent being, a human decision-maker's preference behaviour is generally more complicated than modelled by MAVT.

Another demand of MAVT, which has created some controversy regarding its practicality, is that a gain in effectiveness regarding any criterion must be able to compensate for any loss in effectiveness regarding any other criterion (Antonsson & Otto, 1995). This demand is caused by the application of a weighted sum aggregation function. Such function represents a compensating trade-off between criteria effectiveness, which means that even if an alternative's effectiveness with respect to a particular criterion is zero the overall worth of this alternative can still be non-zero. Otto and Antonsson (1991) call fully compensating trade-offs an 'aggressive' design strategy. Such strategy may not generally be accepted by decision-makers in engineering design (Antonsson & Otto, 1995; Yang & Sen, 1997).

Thurston (1990) applied MAVT to an engineering design decision situation and demonstrated how to avoid an aggressive design strategy. She differentiated criteria into constraints and goals: constraints were binary and had to be satisfied by all feasible alternatives – they were not able to compensate for each other; goals were non-binary, i.e. they could be satisfied to degrees, and they were able to compensate for each other. After all alternatives had been checked for constraint satisfaction, the trade-off between alternatives was then performed on the basis of the goals only. This means that the non-compensating performance characteristics were not included in the effectiveness and overall worth analysis (Thurston, 1990).

Simpson (1996) states that MAVT is not necessarily restricted to the weighted sum. She suggested multiplicative functions or combinations between additive and multiplicative functions. This also helps avoiding aggressive design strategies.

Performance scores

It is necessary to represent the preferences of decision-makers via a set of (effectiveness) scores for the performance of the alternatives regarding the individual criteria (Simpson, 1996). These scores are achieved through the application of an effectiveness function for each individual criterion. To establish such effectiveness function one first needs to define the possible ranges of performances, i.e. below which performance the effectiveness is considered to be zero and above which performance the effectiveness remains unity (Thurston, 1990). Then the actual function is required that maps all possible performances within the defined range onto a numeric scale representing effectiveness. The mapping is a mathematical procedure that requires precise quantitative input in terms of the alternatives' performances. The restriction to quantitative inputs means that even inherently qualitative attributes require precise numerical evaluations (Yang & Sen, 1997).

The establishment of mathematically correct effectiveness functions for individual criteria on the one hand supports MAVT's strong axiomatic foundation. On the other hand they have given rise to much criticism regarding practicality. Buchanan (1994) claims that a practical decision-making method is a trade-off between modelling effort and modelling accuracy. He believes that mathematically correct methods may be cumbersome and unhelpful. Tebay et al. (1984) argue that quantitative methods incur a number of practical measurement problems and thus increased time and cost. For practicality reasons it must be ensured that the decision-making process is not made unnecessarily difficult (Ahn & Dyckhoff, 1997). Ideally, the process should meet both rationality-related and pragmatic requirements. The consideration of rationality-related standards leads to a rational decision while pragmatic requirements refer to the user-friendliness of the method, which is a factor for its actual application. However, both requirement categories are partially in conflict with each other (Ahn & Dyckhoff, 1997).

Regarding input precision, Hazelrigg (1998) states that, strictly speaking, it is not at all possible to know exactly how a particular design will perform until it is built. It has been argued that in particular during the conceptual design phase it is difficult to deliver detailed quantitative input to the decision-making method. This is because in the early design stages the designers are most unsure of the final dimensions and shape, materials and properties and performance of the completed design (Antonsson & Otto, 1995). Höhne (1997) states that a concept is usually described to a large extent in qualitative rather than quantitative terms. Others agree and claim there are very little or even no exact measurements regarding the attributes to be evaluated (Chen & Lee, 1993) and that it may even be difficult to *estimate* expected performances (Thurston & Carnahan, 1992).

Using direct scoring methods, as in table 2.1, does not require the establishment of effectiveness functions. Instead, performance scores may be estimated on an interval scale – theoretically even for inherently qualitative attributes. For example, the riding comfort of a bicycle may be described as being 'good' and a score of '90' on a scale from '1 to 100' may be assigned. The original description is imprecise and qualitative. Yet, the score of '90' is precise. This shows that such numerical scale may actually be rather artificial. Efstathiou (1984) believes that these scores are introduced only for the sake of maintaining an aura of precision. In MAVT and its quantitative derivatives imprecise information must be either ignored or forced to take on a spurious air of precision. This implies that information being in the wrong form does not exist and that greater precision is available than may be justified.

Vadde et al. (1994) conclude that a method which needs precisely defined information is impractical for use in the early stages of design as there is a lack of such information. Methods that require precision regarding inputs are therefore considered more appropriate for later design stages (Matthews et al., 1999).

Weights

In any application, MAVT requires the assignment of weights for the relative importance of criteria (Simpson, 1996). These weights are used by the aggregation function to perform trade-offs between the various criteria. In MAVT the weights are usually directly assigned. This seems simple and straightforward, but Tebay et al. (1984) point out that in practice, assigning weighting factors may actually be rather difficult. It has been argued that direct weight assignments are too abstract for decision-makers and may result in inaccuracies (Zahedi, 1986).

Summary

Using MAVT the decision-makers can clearly see how, via the aggregative model, their beliefs and preferences turn into a suggested overall worth of the alternatives (Simpson, 1996). The method creates not only a relative order of alternatives, but also an absolute measure of overall worth for each alternative.

The main points of criticism against MAVT rest on its apparent failure to act as an appropriately practical decision-making method. The assumptions underlying MAVT seem to be fair and reasonable. However, it seems to be the case that these assumptions are more difficult to satisfy than one might believe. It has been claimed that the general emphasis on numbers and precision may be bewildering to the decision-maker and the sense of accuracy which comes from numerical responses is largely spurious, because numbers are not a natural language for describing phenomena such as human decision-making (Efstathiou, 1984).

2.2.3 The Analytic Hierarchy Process - how it works

The Analytic Hierarchy Process (AHP, Saaty, 1990) is a formal method for determining relationships between discrete alternatives. The essence of this method is pairwise comparisons of criteria and pairwise comparisons of alternatives according to each individual criterion (Žavbi & Duhovnik, 1996). It is based on the idea that a complex issue can be effectively examined if it is hierarchically decomposed into its elements (Azani & Khorramshahgol, 1990). As MAVT, the AHP solves complex decision problems by disaggregating them into a set of sub-problems, solving or reaching consensus on each one, then reassembling them to obtain a solution to the larger decision problem (Weiss & Hari, 1997).

Using the AHP to solve a decision problem can be broken down into four steps (Zahedi, 1986):

1. Establishing a decision hierarchy by breaking down the decision problem into a hierarchy of decision elements. These are the factors that have an influence upon the decision-making situation: overall objective, criteria and alternatives;
2. Assessing the criteria's importance and the alternatives' expected performances through pairwise comparisons;
3. Using the 'eigenvalue technique' to check the consistency of inputs regarding relative weights and performance assessments;
4. Aggregating the relative weights and performance assessments to arrive at an overall ranking of the alternatives.

Step 1: The choice of the decision factors is probably the most creative task in making a decision (Saaty, 1990). In the AHP these factors are arranged in a hierarchical structure, which descends from an overall objective to criteria, to sub-criteria and to alternatives in successive levels. Such a structure offers a 'format' (Hsiao, 1998) for constructing the complex relationships between the various decision factors and it provides an overall view of these relationships. A simple example is given in figure 2.6. Saaty (1990) states that the AHP's hierarchies do not have to be complete. This means, an element in one level does not have to be connected to all elements in the level above, but it may do so.

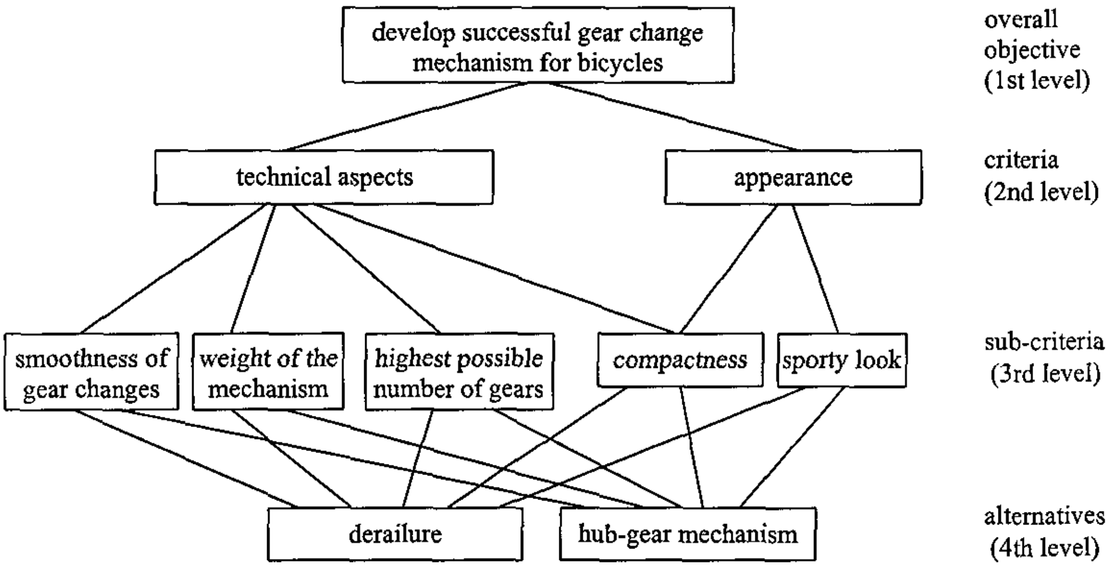


Figure 2.6: Hierarchy of decision elements in the AHP (example).

Step 2: Once the hierarchy is established, the decision-maker is required to create matrices containing pairwise comparisons of same-level-factors. This is done for each level in the hierarchy below the overall objective. A factor at a higher level is said to be a governing factor for those at the next lower level (Khorramshahgol & Moustakis, 1988). Factors at each level are compared to one another regarding the significance or importance of their effect on the governing factor. For example, in the bicycle gear-changing decision, as structured in figure 2.6, the technical aspects' effect upon a successful gear-change mechanism may be three times more important than the effect of appearance. This can be expressed by a very simple matrix, as shown in table 2.2.

The alternatives, placed at the bottom level of the hierarchy, are compared regarding their effectiveness for each criterion in the level directly above, thus creating a number of alternative-effectiveness-comparison matrices, each referring to one criterion in the next higher level.

Step 3: The pairwise comparison matrix as shown in table 2.2 is very simple and should not cause too much difficulty for the decision-maker. However, if more criteria are involved, the decision-maker may have difficulties in being consistent throughout the whole matrix. Consistency here refers to numeric transitivity, for example: if criterion A is two times more important than criterion B and criterion B is two times more important than criterion C, then criterion A should be four times more important than criterion C.

	technical aspects	appearance
technical aspects	1/1	3/1
appearance	1/3	1/1

Table 2.2: Pairwise comparison matrix for AHP.

The AHP's 'eigenvalue technique' (Saaty, 1990) takes the pairwise comparison matrices as an input and computes normalised relative weights of the criteria or relative effectiveness scores of the alternatives with respect to the associated element one level above in the hierarchy. For the example in table 2.2 this would be very easy and the result should be a weight of 3/4 for technical aspects and 1/4 for appearance. In fact, the weights could theoretically be fairly simply calculated from the first row of a comparison matrix of whatever size. The occurring problem, however, is the inconsistency regarding transitivity, as discussed above, apparently inherent (Zahedi, 1986) in most decision-makers' initial comparison matrices. By using the 'eigenvalue technique' these inconsistencies can be detected. The decision-makers then review the pairwise comparisons most likely in an iterative manner. To use this technique it is necessary to complete the entire matrix rather than the first row only.

Step 4: The procedure introduced in the paragraph above calculates normalised weights and performance scores as comparison ratios along the horizontal levels of the hierarchy. The last step is to aggregate these comparison ratios of the various levels along the 'vertical lines' of the hierarchy to produce a statement regarding the relative contribution of each alternative towards the 1st-level overall objective. The AHP's aggregation function is a weighted sum (Antonsson & Otto, 1995). To handle the calculations involved in using the AHP a specific software package, Expert Choice, may be applied.

Application in engineering design

In an experiment, Žavbi and Duhovnik (1996) have applied the AHP to resolve a selection problem in conceptual engineering design. Their issue of concern was the selection of conceptual components for the design of mechanical drive units.

Žavbi and Duhovnik (1996) came to the conclusion that the AHP's two main characteristics, i.e. the pairwise comparisons and the hierarchical criteria structure, are exactly those characteristics that are required for decision-making in conceptual design.

2.2.4 The Analytic Hierarchy Process - characteristics

Underlying assumptions

It is claimed that the AHP's comparisons, as opposed to absolute measurements, describe natural decision-making behaviour in many situations and are therefore considered appealing for practical decision-making (Sen & Yang, 1995; Weiss & Hari, 1997; Pöyhönen & Hämäläinen, 2000). It is assumed that decision-makers are generally capable of expressing their priorities (criteria weights) and preferences (alternatives' performance assessments) through pairwise comparisons. Moreover, as the AHP is a numeric method, it is also assumed that decision-makers are capable of quantifying their pairwise comparisons. Some researchers believe that these assumptions are not always valid. According to Liberatore and Stylianou (1995) it cannot necessarily be assumed that every decision-maker feels generally comfortable with pairwise comparisons and according to Sen and Yang (1995) it is questionable to generally assume that pairwise comparisons are precisely quantifiable.

The AHP is a method for a limited number of criteria and alternatives (Bryson & Mobolurin, 1994). To keep the required input information manageable it was suggested to restrict the number of same-level-factors to a maximum of nine (Zahedi, 1986). Also, as the AHP relates all criteria and alternatives with each other in a rather complex way (hierarchical decomposition and comparison ratios between all factors in each level of the hierarchy), it appears difficult to alter the structure once it has been completed. These aspects suggest that the AHP assumes the criteria-space and the alternatives-space to be static, which is not normally the case in engineering design (Hazelrigg, 1997; Ullman et al., 1997).

Žavbi and Duhovnik (1996) demonstrate how the AHP copes with a dynamic criteria-space and a dynamic alternative-space in terms of including new criteria and alternatives to an existing set already modelled and assessed through the hierarchical structure. Through their description of the required procedures it becomes apparent that adding new criteria or alternatives is indeed possible, but not particularly convenient. This is because most of the AHP procedure (alterations to hierarchical structure, pairwise comparisons, consistency checks, calculation of overall ranking) has to be performed again if a new factor is added. In particular the repetition of the AHP's pairwise comparisons, is said to be difficult and tedious (Ghotb & Warren, 1995).

The AHP is a formal method that has an axiomatic foundation and, similar to MAVT, it applies a weighted sum aggregation function (Antonsson & Otto, 1995). Due to this aggregation function the AHP assumes preference independence (Bryson & Mobolurin, 1994) in the same way as MAVT. This is seen as an important limitation (Liberatore & Stylianou, 1995) for the reasons discussed in section 2.2.2.

Decomposition

For the logical formulation of any complex problem it is often helpful to structure the situation in a hierarchy of related aspects (Ghotb & Warren, 1995). For decision-making these aspects can be translated into criteria, which may be hierarchically decomposed into sub-criteria possibly over a structure of several layers (Sen & Yang, 1995).

Only few decision-making methods actually allow for a hierarchical decomposition of criteria (Antonsson & Otto, 1995). The AHP is one of them. Saaty (1990) states that the AHP hierarchy serves two purposes: It provides an overall view of various relationships in a complex situation and it helps the decision-maker assess whether the criteria in each level are of the same order of magnitude. Hierarchy elements that have a global character can be represented at the higher levels of the hierarchy, whereas others that specifically characterise the problem at hand can be developed in greater depth at the lower levels of the hierarchy.

Apart from improving the problem understanding, hierarchical decompositions are also believed to support communication among a group of decision-makers with little interest in the details of deriving numerical results (Salo & Hämäläinen, 1997). Another aspect of AHP's hierarchical decomposition, pointed out by Žavbi and Duhovnik (1996), is that it enables a recording of knowledge on criteria in particular with respect to their relevance in the overall context

Zahedi (1986) claims the number of levels in an AHP's hierarchy is usually between three and seven: more levels cause a rapid increase in input requirements. It is not clear whether this satisfies decision-making situations in conceptual design.

Pairwise comparisons

One may think that it should be possible to assign importance weights and performance scores directly to the individual criteria and alternatives. Yet, the argument in AHP is that such direct assignments are too abstract for the decision-maker and therefore result in inaccuracies (Zahedi, 1986). Instead, Zahedi (1986) and also Sen and Yang (1995) argue that pairwise comparisons of criteria and alternatives are easy to grasp for decision-makers as they deliver a clear basis for revealing priorities and preferences. Moreover, it is claimed that, in contrast to direct assignments, pairwise comparisons can implicitly also consider cultural context (Žavbi & Duhovnik, 1996) and subjective understanding (Saaty, 1990). On the other hand, it is said that pairwise comparisons may sometimes be difficult and tedious (Ghotb & Warren, 1995). Nevertheless, in conceptual engineering design, in particular for assessing preferences, pairwise comparisons are seen by many researchers as the most practical tool (Chen & Lee, 1993; Žavbi & Duhovnik, 1996; Weiss & Hari, 1997). Their argument is that there are no standard scales or absolute measures available with which the expected performance of a concept may be assessed on an independent, absolute basis.

With respect to criteria importance it is argued that designers cannot necessarily, a priori, specify their priorities; only preliminary estimates; they need to gain insight on how to specify their weight assignments through iteration (Otto & Antonsson, 1991). In AHP the decision-maker first establishes a set of weights through pairwise comparisons and then checks the consistency of these weights by applying the 'eigenvalue technique'. If inconsistencies occur, the decision-makers may reconsider the weighting information, take corrective action and reapply the 'eigenvalue technique'. Thus, the AHP supports iterative weight specification. Some researchers consider this approach as one of AHP's particular advantages (Ghotb & Warren, 1995; Žavbi & Duhovnik, 1996) as it seems to gain consistent weights without requiring an excessive commitment of time (Bryson & Mobolurin, 1994). However, other researchers (Naudé et al., 1997) consider this approach, in particular the application of the 'eigenvalue technique' as one of AHP's particular disadvantages. They claim it is strenuous for the decision-maker and produces spurious large and small weights. It has also been argued that the AHP's weight assignment approach creates much data that is actually redundant, which is counterproductive as it leads to inconsistency in the first place (Sen & Yang, 1995).

Sensitivity of comparisons

When using the AHP's pairwise comparison approach for assessing weights and performances, the decision-maker needs to quantify these comparisons on a discrete ratio scale from '1' (equal) via '3', '5' and '7' to '9' (strongly prioritised/preferred). The general applicability and meaning of this ratio scale for assessing weights and performances has been debated. Chen and Lee (1993) claim that the only information available for assessing the expected performance of alternatives, at least in conceptual design, is whether or not the one alternative's performance is preferred over the other alternative's performance. Ghotb and Warren (1995) think that the AHP's '1 to 9' scale is not sensitive enough, as it does not translate marginal differences. Salo and Hämäläinen (1997) believe, generally, the '1 to 9' scale severely restricts the range and distribution of possible criteria priorities and performance preferences. Yet, it is also claimed that due to psychological limitations the AHP's '1 to 9' scale does not need to be any more sensitive as it is (Jerčić & Bajić, 1993). It has also been suggested that no fixed numerical scale should be used as a standard tool for every situation because this can lead to biased numerical estimates of priorities and preferences due to the specific characteristics of particular situations and decision-makers (Pöyhönen et al., 1997).

Linguistic comparisons

An alternative way for expressing priorities and preferences through AHP's pairwise comparison approach is using a scale of linguistic expressions rather than the numeric scale (Ghotb & Warren, 1995). From the practitioners' point of view the possibility of using linguistic expressions is considered a very attractive feature of the AHP (Pöyhönen et al., 1997). Using AHP's linguistic scale the decision-maker has the option of expressing priorities or preferences as 'equal', 'weakly prioritised/preferred', 'prioritised/preferred', 'strongly prioritised/preferred' or 'absolutely prioritised/preferred' (Zahedi, 1986).

A probable reason for the apparent popularity of linguistic expressions for priority and preference assessments could be that decision-makers implicitly assume that linguistic responses allow some ambiguity or robustness and that a method using linguistic comparisons automatically takes into account ambiguity when deriving the results; such principles are, however, not included in the AHP (Pöyhönen et al., 1997). To process the linguistic expressions, AHP translates them into the numeric integer ratios '1', '3', '5', '7' or '9' (Zahedi, 1986), which ignore all ambiguity inherent in linguistic expressions. Experiments have shown that different decision-makers have different interpretations regarding the same linguistic expression (Pöyhönen et al., 1997). Therefore, a standard mapping of linguistic expressions onto crisp numbers, no matter who is responding and in what context, must be regarded with due caution (Salo & Hämäläinen, 1997).

Summary

The AHP's particular appeal is based on the pairwise comparisons for priority and preference assessments (Salo & Hämäläinen, 1997). This is said to be especially convenient in situations where it is difficult or impossible to measure performances (Žavbi & Duhovnik, 1996). Much credit is also given for its hierarchical criteria decomposition (Barzilai, 1997).

Criticism of the AHP addresses the practicality of its approach to check consistency in the pairwise comparison ratios as well as the actual ratio scales applied. It is claimed that, due to its excessive input demands, the technique for consistency checks actually furthers inconsistency (Sen & Yang, 1995) apart from making the provision of input strenuous (Naudé et al., 1997). Input may be provided either by using a numeric or by using a linguistic ratio scale. The numeric ratio scale assumes that comparisons are precisely quantifiable, which is seen questionable (Sen & Yang, 1995). The linguistic ratio scale needs to be regarded with due caution as it does not actually process any ambiguity inherent in linguistic expressions (Pöyhönen et al., 1997). This is because eventually the linguistic expressions are mapped by the AHP onto its standard scale, which consists of numeric integer scores. It is debated whether such a standard scale may be applied for all situations and whether it allows for appropriately quantifying different decision-makers' interpretations of linguistic expressions (Salo & Hämäläinen, 1997).

Another criticism that may be raised against the AHP regards its rigidity; it seems to assume static decision-spaces, which is, at least in engineering design, not normally the case (Hazelrigg, 1997; Ullman et al., 1997). Apart from this, the AHP assumes preference independence (Bryson & Mobolurin, 1994) in the same way as MAVT. This is seen as an important limitation (Liberatore & Stylianou, 1995).

In contrast to MAVT, the AHP only creates a relative order of the alternatives, but no absolute measure of overall worth. Yet, AHP's relative order is quantitative, which means that the differences regarding overall worth are indicated.

2.2.5 Outranking – how it works

Outranking is based on the establishment of outranking relations between alternatives. An outranking relation is a binary relation which compares the arguments for and against the hypothesis 'alternative A is at least as good as alternative B' (in the literature usually denoted as 'ASB'), given what is known about the decision-maker's preferences (Simpson, 1996).

Outranking relations are either restricted or comprehensive. A restricted outranking relation states the level of preference of one alternative over another alternative (pairwise comparison) with respect to one criterion only. This comparison is based on the alternatives' anticipated performance. A comprehensive outranking relation states the preference of one alternative over another alternative with respect to all criteria.

Let us assume that A and B are alternative solutions for a particular problem and that $g_j(A)$ and $g_j(B)$ are the performances of A and B with respect to a maximising (high is preferred) criterion j. Such performance needs to be expressed by a real number even if it reflects a qualitative assessment (Roy, 1991).

The restricted outranking relation AS_jB holds if the performances $g_j(A)$ and $g_j(B)$ give a strong enough argument for considering alternative A as being at least as good as alternative B, with respect to criterion j. 'At least as good as' must be understood as 'not worse than', which may mean either (i) $g_j(A)$ is indifferent to $g_j(B)$, AI_jB; (ii) $g_j(A)$ is weakly preferred over $g_j(B)$, AQ_jB; or (iii) $g_j(A)$ is strictly preferred over $g_j(B)$, AP_jB. The distinction between these cases can be found by first calculating and then comparing the difference between $g_j(A)$ and $g_j(B)$ to some 'thresholds': indifference threshold q_j and preference threshold p_j. These thresholds allow for the comparison of performances in an imprecise manner (Simpson, 1996).

The indifference threshold q_j, which determines indifference between $g_j(A)$ and $g_j(B)$, is defined such that (Roy, 1991):

$$AI_jB \text{ iff } g_j(A) - g_j(B) \leq q_j$$

$$\text{Thus: } AS_jB \text{ iff } g_j(A) \geq g_j(B) - q_j.$$

The preference threshold p_j, which determines strict preference of $g_j(A)$ over $g_j(B)$, is defined such that (Roy, 1991):

$$AP_jB \text{ iff } g_j(A) > g_j(B) + p_j.$$

Weak preference of $g_j(A)$ over $g_j(B)$ is characterised by the relation AQ_jB, which holds in case (Roy, 1991):

$$g_j(A) - p_j \leq g_j(B) < g_j(A) - q_j \qquad (p_j > q_j).$$

A graphical representation of these thresholds and relations, based on Simpson (1996), is shown in figure 2.7.

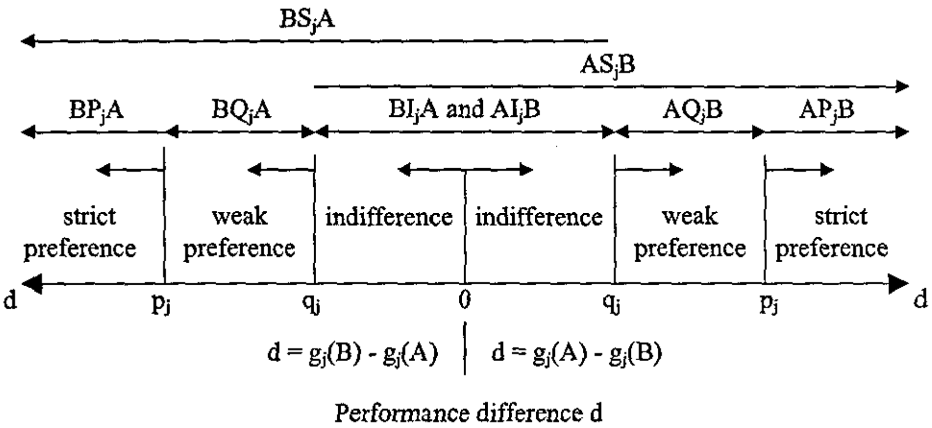


Figure 2.7: Graphical representation of outranking thresholds.

The aggregation of the restricted outranking relations into comprehensive outranking relations is performed differently by the various outranking methods. Two well-known types of these methods (Diakoulaki & Koumoutsos, 1991) are Electre (ELimination Et Choice Translating REality) and Promethee (Preference Ranking Organisation METHod for Enrichment Evaluations).

Electre

The comprehensive outranking relation ASB holds if the entire set of restricted outranking relations give a strong enough argument for considering alternative A as being at least as good as alternative B. The basic concepts of the validation of the outranking relation ASB in Electre (Roy, 1991) are those of concordance and discordance, which are based on the application of thresholds, and on criteria importance coefficients, i.e. weights.

A particular criterion j is in concordance with the assertion ASB if AS$_j$B holds, i.e. if the performance $g_j(A)$ is at least as good as the performance $g_j(B)$. A particular criterion j is in discordance with the assertion ASB only if the performance $g_j(B)$ is strictly preferred to the performance $g_j(A)$. A particular criterion j may be neither in concordance nor in discordance with the assertion ASB. This happens if $g_j(B)$ is weakly preferred over $g_j(A)$. Thus, the entire set of criteria can be partitioned into three subsets: the criteria which are in concordance with the assertion ASB, the criteria which are in discordance with the assertion ASB, and the criteria which are neither in concordance nor in discordance with the assertion ASB.

The concordance index characterises the strength of the positive arguments able to validate the assertion ASB. The strongest arguments come from those criteria that belong to the concordance partition. Their strength is measured by the ratio between the sum of their importance weights and the sum of all criteria importance weights. Some weaker positive arguments can also come from criteria which belong to the 'neither concordance nor discordance' partition. This is because the decision-maker hesitates to declare that $g_j(B)$ is strictly preferred over $g_j(A)$. The strength of these criteria is measured by the ratio between the sum of their reduced importance weights and the sum of all criteria importance weights. The reduction of the importance weights depends on how close to indifference the weak preference of $g_j(B)$ over $g_j(A)$ is. The closer the difference $g_j(B)$ - $g_j(A)$ to the indifference threshold the stronger is the support for the assertion ASB and thus, the less is the reduction of the importance weight. Roy (1991) describes in detail how to calculate reduced importance weights.

The numeric value of the concordance index is used by Electre for the validation of the assertion that 'alternative A is at least as good as alternative B', ASB. A high concordance index supports this assertion and a low concordance index opposes this assertion. The meaning of 'high' and 'low' is determined by a limiting value, which has to be elicited from the decision-maker.

The discordance index characterises the strength of the negative arguments able to validate the assertion ASB. These arguments come from criteria that belong to the discordance partition. Some of these criteria might be associated with restricted outranking relations which not only represent strict preferences of $g_j(B)$ over $g_j(A)$, but in addition the difference $g_j(B)$ - $g_j(A)$ exceeds a veto threshold v_j. The veto threshold indicates the point at which an alternative is so outperformed on one criterion that the hypothesis 'A is at least as good as B' does not hold overall, regardless of all other criteria and the value of the concordance index (Simpson, 1996).

It may occur that two alternatives cannot be compared. This is the case if neither the hypothesis 'alternative A is at least as good as alternative B' nor the hypothesis 'alternative B is as least as good as alternative A' have strong enough arguments (values of concordance and discordance indices) for their support. This may occur when one alternative is fairly better on a subset of criteria and the other one is fairly better on another subset other criteria with similar added weights (Wolters & Mareschal, 1995). If the alternatives A and B are not comparable, they cannot be ranked.

Promethee

Promethee (Brans et al., 1986) is based on the same foundation as Electre. Yet, the aggregation of restricted outranking relations into comprehensive outranking relations is carried out in a different way. Promethee excludes incomparabilities between alternatives and thus provides complete rankings across all alternatives (Diakoulaki & Koumoutsos, 1991).

The Promethee rankings are obtained by considering 'flows' in outranking graphs. An example of such a graph is shown in figure 2.8. Between the alternatives A and B there are two flows, one from alternative A to alternative B and another one vice versa. The flows are determined by the preference indices, $\pi(A,B)$ and $\pi(B,A)$. These indices represent the decision-maker's preference intensity of one alternative over the other, when considering all criteria (comprehensive outranking relation). The preference index is calculated in a similar fashion as the concordance index discussed above. However, it is not necessary to establish a 'limiting value' for the concordance index as required in Electre. It is not necessary either to take any discordance into account (De Keyser & Peeters, 1994).

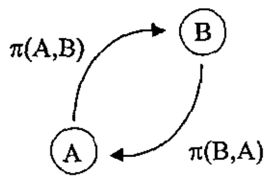


Figure 2.8: Binary outranking graph.

If there are more than two alternatives involved, then each node in an outranking graph would be connected to a number of other nodes, i.e. alternatives, as shown in figure 2.9. The leaving flow of an alternative A, $\Phi^{+}(A)$, is then defined as the sum of the values of the flows leaving node A and therefore provides a measure of the outranking character of A over all other alternatives. The entering flow of an alternative A, $\Phi^{-}(A)$, is defined as the sum of the values of the flows entering node A and therefore provides a measure of the outranked character of alternative A with respect to all other alternatives. The difference between leaving flow and entering flow of an alternative A is called its net flow, $\Phi(A) = \Phi^{+}(A) - \Phi^{-}(A)$. Alternative A outranks alternative B if $\Phi(A) > \Phi(B)$ and alternative A is indifferent to alternative B if $\Phi(A) = \Phi(B)$. This net flow forces all alternatives to be comparable and thus enables the establishment of a complete ranking order.

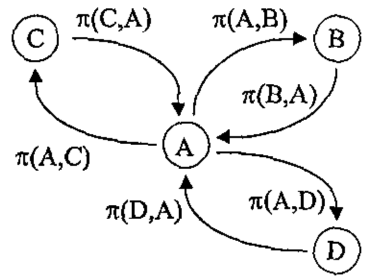


Figure 2.9: Multiple outranking graph.

Application in engineering design

The basic outranking principle has also been applied to methods specifically developed for decision-making in engineering design. There are two very 'traditional' methods (Otto & Antonsson, 1991), which are decision-making with Quality Function Deployment (QFD) and the method of controlled convergence.

Quality Function Deployment

QFD (Hauser & Clausing, 1988) aims at linking abstract customer requirements to the specific decision-making processes within a company (Wright, 1998). It aims at ensuring that the customers' needs are considered throughout the product development process. QFD is based upon a system of interlinked charts, which will be briefly discussed in the following.

The first QFD chart relates the needs stated by the customers (the desired performance variables) with the directly measurable engineering characteristics (the design variables). A number of engineering characteristics may be linked to a single customer need. The customer needs address high-level (abstract) criteria for the evaluation of alternative designs, whereas the engineering characteristics address low-level (specific) criteria for the evaluation of alternative designs. The QFD chart incorporates symbols, which allow for expressing three levels of relationships between customer needs and engineering characteristics: strong, medium, and weak. By using these symbols and relative importance ratings (RIR) of customer needs an indication of the relative importance of the engineering characteristics can be achieved. The chart also allows for representing influential relationships between engineering characteristics. A QFD chart example, simplified from Wright (1998), for a car door is shown in figure 2.10. The left-hand side of the chart is a prioritised list of customer needs. The upper part of the chart is a list of related engineering characteristics and the triangle on the very top depicts influences between engineering characteristics. The matrix in the middle of the chart shows the relations between customer needs and engineering characteristics.

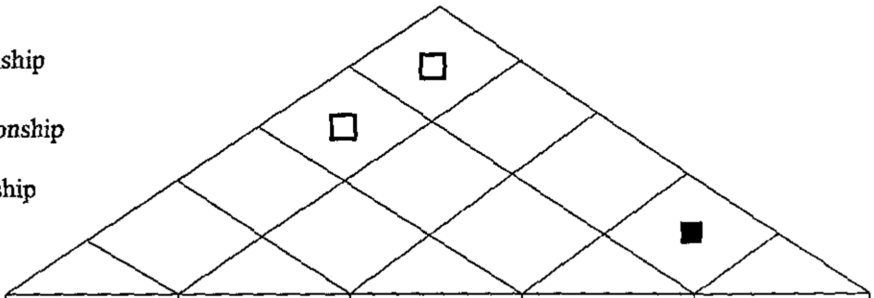
The second QFD chart relates engineering characteristics to parts' characteristics. The third chart relates the parts' characteristics to key process operations and a fourth chart relates the key process operations to production requirements. Thus, the initially stated customer requirements can be related to all stages of the product development process. Having said this, if applied in industrial practice, companies mainly limit the use of QFD to the first, or the first and second charts because of either insufficient

Key:

■ Strong relationship

□ Medium relationship

• Weak relationship



engineering characteristics / customer needs	RIR	closing effort (Nm)	inside handle opening effort (N)	outside handle opening effort (N)	static force to hold open (N)	dynamic force to hold open (Nm)
easy to close from outside	7				•	•
easy to close from inside	5	■			□	□
easy to open	2		■	■		•
stays open on hill	5				■	□

Figure 2.10: QFD chart for car door.

information referring to the remaining charts or because of time overheads involved in their construction (Wright, 1998).

Decision-Making with Quality Function Deployment

Hales (1995) describes in detail how QFD may be directly applied as a decision-making method: At first a prioritisation matrix is constructed which aims at enabling the decision-maker to translate the prioritised list of customer needs into a prioritised list of engineering characteristics, or measures of alternative quality. This matrix looks like the chart in figure 2.10. However, the levels of relationship are associated with weights, such as strong equals 9, medium equals 6, and weak equals 1. The relationship weights are multiplied with the row weights, which prioritise the customer needs. The products are added up for each column and the results, normalised over all characteristics, prioritise, i.e. weight, the list of engineering characteristics, which are the low-level evaluation criteria.

The second step is the construction of a selection matrix, which consists of the prioritised list of engineering characteristics on the left-hand side and a list of alternatives on the upper part of the matrix. The engineering characteristics are used as the low-level evaluation criteria for the actual selection process. This process uses similar principles as the outranking methods. These principles are: pairwise comparisons of alternatives with respect to individual, weighted criteria; assessing whether the one alternative's performance is preferred to or indifferent to the other alternative's performance (no weak preferences are considered); establishing concordance and discordance indices (Hales does not actually use exactly these terms, but describes the principle) for all alternatives; and applying these indices for generating an overall ranking of the alternatives.

In decision-making with QFD, as described by Hales (1995), the alternatives are not all compared to each other, but each alternative is compared to a datum only. This datum may be an already existing design, e.g. the current product's concept, or any other suitable reference chosen by the decision-maker (Weiss & Hari, 1997).

If, with respect to some specific criterion, the alternative's performance is preferred to the datum's performance, a plus is put into the matrix square that relates this alternative with the criterion in question. If the alternative's performance is considered indifferent to the datum's performance, an 's' is put into the matrix. If the datum's performance is preferred to the alternative's performance a minus is put into the matrix.

The concordance index is calculated by adding up the weights of all criteria upon which the alternative's performance is preferred over the datum's performance. The discordance index is calculated by adding up the weights of all criteria upon which the datum's performance is preferred over the alternative's performance. Weights of criteria upon which the alternative's performance is considered indifferent to the datum's performance are ignored for the calculation of concordance and discordance indices. Finally, a comprehensive 'outranking' statement is found by subtracting the discordance index from the concordance index. An abstract example is given in table 2.3.

Engineering characteristics (*criteria*)	Weights	Alternatives			
		A	B	C	D
X	5	+	-	-	DATUM
Y	3	+	+	-	
Z	2	s	+	s	
Concordance		8	5	0	
Discordance		0	5	8	
Σ		8	0	-8	

Table 2.3: QFD selection matrix.

Using the above procedure does not actually allow for establishing an overall ranking across all alternatives. This is because not all alternatives are compared to each other and the alternative-datum comparisons are based on an ordinal scale only. This means, it is expressed whether or not an alternative's performance is preferred to the datum's performance, but not by how much. Hales (1995) suggests that this problem may be overcome by repeating the procedure with different alternatives used as datum.

Pugh's method of controlled convergence

The method of controlled convergence (Pugh, 1990), or also called datum method (Roozenburg & Eekels, 1995), is considered a 'classic' in the area of conceptual engineering design (Cziulik & Driscoll, 1997). It is based on a selection matrix similar to the one described above. Yet, no criteria weights are established. For calculating the concordance and discordance indices all criteria are given the factor '1'. This means that an alternative's concordance index is represented by the number of plusses, its discordance index is represented by the number of minuses. The assessment of an alternative's 'overall profile' is based on these indices. A high concordance index, i.e. many pluses, denote a strong alternative and a high discordance index, i.e. many minuses, denote a weak alternative. If most of the alternatives tend to have more pluses than minuses this indicates that the chosen datum is a weak alternative for itself. A selection matrix as for the method of controlled convergence is shown in table 2.4.

Engineering characteristics (criteria)	Alternatives				
	A	B	C	D	E
U	+	-	+	+	DATUM
V	+	s	+	+	
W	-	+	-	+	
Y	-	+	+	S	
X	-	-	+	+	
Z	-	-	s	S	
Concordance (Σ+)	2	2	4	4	
Discordance (Σ-)	4	3	1	0	

Table 2.4: Selection matrix for the method of controlled convergence.

The comparison described above is repeated a few times with different alternatives serving as a datum. Pugh (1990) suggests that by doing this one will eventually converge towards one or a small number of strong alternative concepts and it will become clear which alternatives are the weak ones. No formal aggregation function is used (Roozenburg & Eekels, 1995). Therefore, it can be said that Pugh's method supports systematic, but informal decision-making (Antonsson & Otto, 1995).

2.2.6 Outranking – characteristics

Underlying assumptions

The outranking methods start from the assumption that a decision-maker can only approximate comparisons of the alternatives' performances (Simpson, 1996). According to Roy (1991) this is due to the existence of uncertainty in a human decision-maker's mind, of half-held beliefs and of imprecision in the determination of the alternatives' performances. Such uncertainty and imprecision is accounted for through the introduction of zones or ranges of indifference and preference between performances (Simpson, 1996) as opposed to precise quantifications. The ability to cope with decision-makers who hesitate when they are asked to precisely state their preferences is seen as a major advantage of the outranking methods over approaches like MAVT and the AHP (Roy, 1991).

In contrast to MAVT and the AHP, outranking methods have no axiomatic foundation; instead they are based on parameters (limits and indices) and an intuitive decision algorithm (Simpson, 1996). It is argued that due to the missing axiomatic foundation, outranking relations may not necessarily be transitive (De Keyser & Peters, 1996; Roy, 1991). The axiomatic foundation of MAVT and the AHP assumes preference independence. It does not seem to be absolutely clear whether or not this assumption must hold for the outranking methods as well (Simpson, 1996).

The outranking methods assume that a decision-maker is able to express their preferences as pairwise comparisons (De Keyser & Peters, 1996). In contrast to AHP, these comparisons are not quantified, but stated in qualitative terms. Such qualitative comparisons imply the assumption that little differences in the compared alternatives' performances are practically irrelevant and that performance differences need to be of some magnitude before they have a bearing on preferences (Simpson, 1996).

A decision-maker using an outranking method is only required to provide relatively little input information (Diakoulaki & Koumoutsos, 1991). This is mainly due to the qualitative character of the pairwise comparisons. On the one hand this may be seen as a strength, because it simplifies the decision-making process, but on the other hand this may also be seen as a weakness because it leads to a lack of quantitative differentiation regarding the involved alternatives (Diakoulaki & Koumoutsos, 1991). Outranking methods are limited to ordinal differentiations (Simpson, 1996).

Having said above that relatively little input information is required, outranking still assumes that a decision-maker is able to determine some parameters. According to the particular method used these are thresholds on each criterion, limits for concordance and discordance indices and precise criteria weights. The elicitation of meaningful and justifiable values for these parameters is not a trivial exercise (Simpson, 1996). It is believed, decision-makers cannot usually set correctly their exact numerical values for the parameters involved (Brans et al., 1986), neither for the thresholds and indices, nor for the criteria weights (Roy, 1991). The decision-maker is assumed to be a 'rational economic person' (Simpson, 1996), which means they are able to compare the alternatives performances and set the parameters involved by using 'common sense' instead of mathematically rigorous functions.

Restricted outranking relations

Establishing a restricted outranking relation basically involves comparing two alternatives' performances with respect to some specific criterion and determining whether one of the performances is preferred over the other one or whether they are considered indifferent. Preference and indifference hold over intervals of actual performance differences. These intervals are determined through thresholds. Three thresholds were introduced in section 2.2.5: the indifference threshold, the preference threshold and the veto threshold.

Various threshold concepts for restricted outranking relations are discussed by Huylenbroeck (1995) and by Brans et al. (1986). The concept using an indifference threshold and a preference threshold as shown in figure 2.7 is called 'level criterion' (Brans et al., 1986) or 'pseudo criterion' (Roy, 1991). A more sophisticated concept, the multi-level criterion, using two additional thresholds is discussed by Huylenbroeck (1995). These two additional thresholds extend the level criterion's three comparison intervals to altogether five intervals. On the one hand such extension allows for a more sensitive comparison of the alternatives' performances, but on the other hand it requires the decision-maker to set another two thresholds. Obviously, it needs to be considered whether these preference differentiations are actually meaningful for the particular comparison situation at hand (De Keyser & Peters, 1996).

One of the simpler concepts is the 'quasi-criterion' (Brans et al., 1986; Roy, 1991), also called '0-1 criterion with indifference area' (Huylenbroeck, 1995). This concept does not involve a preference threshold. Thus, there are only two comparison intervals, indifference and preference, separated by the indifference threshold.

The quasi-criterion is used for decision-making with QFD, as described by Hales (1995), and for Pugh's method of controlled convergence. Although these methods do require the decision-maker to apply an indifference threshold, they do not actually require the decision-maker to be explicit about this threshold. The application of such a method is seen as being particularly easy and straightforward (Hales, 1995) as the decision-maker is only confronted with very simple choice situations. Yet, Matthews et al. (1999) argue that such a method suffers from a considerable degree of subjectivity by the decision-maker. These methods are seen to be applicable for supporting judgements about qualitative information (Ullman & D'Ambrosio, 1995).

Comprehensive outranking relations

For establishing comprehensive outranking relations with Electre it is necessary to set further parameters: limits for concordance and discordance indices. This is not necessary if Promethee, QFD, or the method of controlled convergence are used.

The establishment of comprehensive outranking relations through limiting values for concordance and discordance indices as in Electre may lead to incomparability of alternatives. Some researchers consider this characteristic as a particular strength of Electre, as information about incomparability is seen to be very valuable for decision-making (Brans et al., 1986; Wolters & Mareschal, 1995). On the other hand, the practical usefulness of such information is questioned, as it does not help for making a choice between a number of available alternatives (Simpson, 1996).

Enforcing comparability between all alternatives, although probably being less realistic than allowing for incomparability, is seen as the more concrete approach, making it easier to actually resolve a decision-making problem (Brans et al., 1986; Diakoulaki & Koumoutsos, 1991). The particular merit of generally enforcing comparability is the possibility to establish complete rankings, in contrast to partial rankings, across all alternatives. However, even if a ranking is complete it may still be questionable in terms of its practical usefulness. De Keyser and Peters (1996) demonstrate that complete rankings, established with Promethee, may be very unstable. This means that the relative position of alternatives in the ranking order can change by adding or deleting an alternative. De Keyser and Peters (1996) show that even adding or deleting alternatives which are totally dominated or equal to another alternative may change the ranking order to such an extent that a different decision (the highest ranked alternative) may be implied. This is seen as a questionable characteristic (De Keyser & Peters, 1996).

Most of the outranking methods calculate comprehensive outranking relations by using criteria importance weights, which need to be precisely defined.

Decision-making with QFD offers a mechanism (the prioritisation matrix) for building relations between the decision criteria and the customer requirements (Hales, 1995). It is admitted, however, that this mechanism does not lead to very accurate criteria weights (Hales, 1995) and may be tedious, if not impractical (Geng et al., 1996).

Franceschini and Rosetto (1995) claim, the conversation of the customers' relative importance ratings (RIR) into weights for the engineering characteristics, i.e. the low-level evaluation criteria, is somewhat arbitrary. They argue that the customers' relative importance ratings will in fact be an ordinal ranking which is then transformed through the relationship matrix and the 1-6-9 factors into a cardinal weighting for the engineering characteristics. Such an ordinal-cardinal transformation is questionable. It assigns an arbitrary precision to the customers' assessments and it imposes an 'element of absolute truth' on the customers' evaluations pretending more informative content than has in fact been expressed (Franceschini & Rosetto, 1995).

The method of controlled convergence does not use any aggregation function for establishing comprehensive outranking relations. Therefore, the decision-maker is not required to assign criteria weights. This method results in a rather abstract satisfaction expectation for each alternative (Ullman & D'Ambrosio, 1995). The method of controlled convergence cannot aim at the straightforward selection of a highest ranked alternative, but it results in suggestions for improving particular aspects of the involved alternatives and as such supports design synthesis (Weiss & Hari, 1997).

Summary

The outranking methods are based on qualitative pairwise comparisons. When using these methods for evaluating the alternatives' performances the decision-maker is only required to provide relatively little input information (Diakoulaki & Koumoutsos, 1991). This is seen as a particular strength, because it simplifies the decision-making process. Another aspect which is considered appealing is that these methods are able to cope with decision-makers who hesitate when they are asked to precisely state their preferences (Roy, 1991). It is only assumed that a decision-maker can approximate comparisons of the alternatives' performances (Simpson, 1996).

The outranking methods have no axiomatic foundation; instead they are based on parameters (limits and indices) and an intuitive decision algorithm (Simpson, 1996). Because of this, it cannot be assumed that outranking relations are necessarily transitive (De Keyser & Peters, 1996; Roy, 1991) and it is not absolutely clear whether or not preference independence must hold (Simpson, 1996).

Problems may occur when it comes to determining the various outranking parameters. However, this depends on the particular method used. Electre requires the decision-maker to set thresholds for restricted outranking relations, limits for the concordance and discordance indices used to establish comprehensive outranking relations and precise criteria weights are also needed. On the other extreme, the method of controlled convergence does not need any parameter to be explicitly expressed.

The main shortcoming of the outranking methods is said to be the lack of quantitative differentiation between alternatives (Diakoulaki & Koumoutsos, 1991).

In contrast to MAVT, the outranking methods only create a relative order of the alternatives, but no absolute measure of overall worth. This is similar to the AHP. The difference to the AHP is that a relative order established through outranking has a mere ordinal nature (Simpson, 1996). This means, that one can say 'alternative A is preferred over alternative B', but one cannot say by 'how much' as in the AHP (Diakoulaki & Koumoutsos, 1991).

2.3 Methods for modelling uncertainty

The methods for multi-criteria decision-making as discussed within the previous sections do not model uncertain information. MAVT and the AHP assume all information is known with certainty. The outranking methods acknowledge uncertainty, but do not actually model it.

Uncertainty may be attached upon much of the information available in practical decision-making situations. Decision-making in conceptual design does not seem to be an exception: Hazzelrigg (1998) argues that it is not possible to know with certainty how a particular design will perform until it is actually built, but it cannot be built until its underlying concept is selected. This would suggest that decision-making in conceptual design always takes place under conditions of uncertainty. Uncertainties may not only be attached to performance information, but also to information regarding the importance of criteria (Wolters & Mareschal, 1995).

A number of formal methods have been developed that allow for modelling uncertain information. These methods have been integrated into decision-making methods. Thus, the overall worth measures used to rank order the alternatives are then also valid under conditions of uncertainty (Hazzelrigg, 1998). This, in turn, helps to produce robust designs (Antonsson & Otto, 1995) by preventing uncertain information from being ignored for the systematic analysis of decision-making situations. Prevailing methods for representing uncertainty are (Antonsson & Otto, 1995): Probability theory, Bayesian inferencing, Dempster-Shafer theory, and Fuzzy set theory. These methods will be discussed in the following sections.

2.3.1 Probability theory

Probability theory is a mathematical method that is useful for discovering and investigating the regular features of random events (Mizrahi & Sullivan, 1993). Randomness is linked with the occurrence of chance phenomena. These phenomena may not only produce different outcomes each time they are observed but also, the occurrence of a particular outcome may not be predictable. However, there is a long-range behaviour, which is known as statistical regularity.

A typical example of such a phenomenon is throwing a dice. In a series of throws with a fair dice, each throw results in one of the numbers 1, 2, 3, 4, 5, or 6. Since it is not predictable which number will appear, the result of throwing a dice is a random event. Yet, if the dice is fair and a sufficiently great number of events is observed it will be noticed that the different faces of the dice occur at an approximately equal number of times. Thus, it is reasonable to say that each face, i.e. number, is equally likely to occur. As a result, one may assign the probability of 1/6 for obtaining a particular number at a particular throw.

In studying probability one is concerned with experiments and their outcomes. For analysing and predicting the outcomes of such experiments a probability model can be built. Building a probability model requires two steps:

1. List all possible outcomes of the experiment. The set of all possible outcomes of the experiment together constitute the sample space. In the example above, an outcome is the appearance of a particular face and the sample space consists of the six faces of the dice.
2. Assign to each outcome e a probability P(e) so that

- $P(e) \neq 0$, in the example above $P(e) = 1/6$ for all e;
- The sum of all probabilities over the entire sample space equals 1, in the example above $6 \cdot 1/6 = 1$.

Probabilities can be assigned to events and simple events. An event is a subset of the sample space. If an event only has one element, i.e. consists of only one outcome, it is called a simple event. Every event that is not a simple event can be written as the union of simple events. For example: If we throw the dice two times, a possible event is that we have the occurrence of a 5 and a 6. This event is the union of the simple events (a) first a 5 and then a 6 and (b) first a 6 and then a 5. The probabilities of the occurrence of the simple events are the products of the probability of obtaining a particular face at a single throw, $1/6 \cdot 1/6 = 1/36$. The probability of the occurrence of the event is the sum of the probabilities of the occurrence of the simple events, $1/36 + 1/36 = 1/18$. This structure can be delineated in a probability tree as in figure 2.11.

A probability tree can also be seen as a decision tree. Each circle, as depicted in figure 2.11, represents a decision and the lines emerging from the circles represent the possible outcomes of these decisions. The probability values attached to the lines represent the likelihood of each outcome to occur. Once an outcome has occurred a new decision is made which in turn may result in a range of outcomes. By means of a decision tree whole strategy scenarios can be probabilistically investigated.

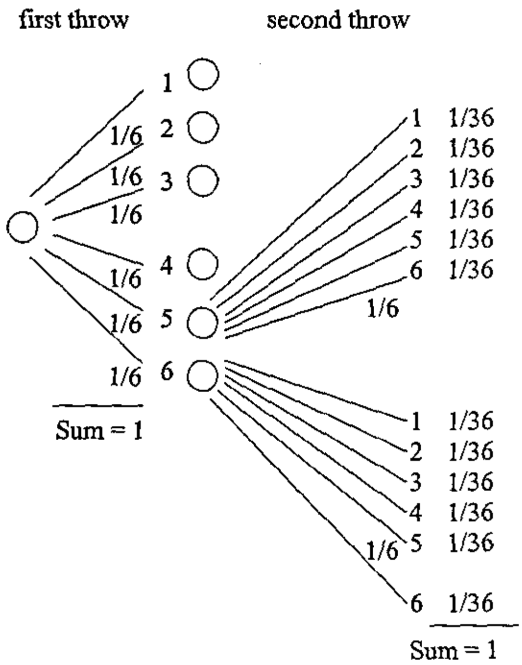


Figure 2.11: Probability tree.

2.3.2 Application of probability theory – how it works

An application of probability theory in decision-making is multi-attribute utility theory (Keeney & Raiffa, 1976), which allows the decision-maker to express uncertainty about the expected performance of alternatives. Multi-attribute utility theory (MAUT) is based on MAVT. MAUT's principle for gaining a utility for an uncertain performance is to state a utility for an alternative's performance level that is certain versus one that will occur with some probability (Thurston & Carnahan, 1992).

For establishing the utility of an uncertain performance the decision-maker initially needs to respond towards two questions (Sage, 1995): First, what are possible performances of the alternative under consideration (with respect to some criterion) and what is their effectiveness? Second, what is the probability of the occurrence of the various performances? Responding towards these questions leads to the establishment of a 'lottery', L.

The effectiveness of performances is determined exactly in the same way as in MAVT, i.e. using an effectiveness function. An (exaggerated) example is: a particular car alternative may result in two possible outcomes with respect to the criterion 'fuel economy'. With reference to the example effectiveness function in figure 2.3, performance 1 may be a fuel economy of 70mpg, which represents an effectiveness of 0.8, with a probability of 0.3 and outcome 2 may be a fuel economy of 30mpg, which represents an effectiveness of 0.35, with a probability of 0.7. The problem is now to determine a utility for this particular car alternative with respect to fuel economy. This can be done by using a certainty equivalent.

A certainty equivalent of a lottery L is a performance X such that the decision-maker is indifferent between L and the performance X for certain (Keeney & Raiffa, 1976). The effectiveness of that performance X, as determined by the criterion's effectiveness function, is then considered as this alternative's utility with respect to the criterion in question. The alternatives' 'overall worth' can be established, as in MAVT, by using an aggregation function.

2.3.3 Application of probability theory – characteristics

Assumptions

MAUT does not model uncertainty with respect to the criteria importance weights (Thurston & Carnahan, 1992). This means that it is assumed these weights can be precisely determined with certainty. Yet, it has been discussed in previous sections that assigning precise criteria weights is not generally considered a straightforward exercise. What can be modelled with MAUT is uncertainty with respect to the alternatives' expected performances (Antonsson & Otto, 1995).

As MAUT is entirely based on probability theory it is assumed that the involved uncertainties are of a random nature. Some researchers believe that, with respect to performance expectations in engineering design, this is a fair assumption (Lacksonen, 1995; Hazzelrigg, 1998). Yet, other researchers believe that this assumption is not generally valid, as in many cases the involved uncertainty is not random in nature (Efstathiou, 1984), but rather due to vagueness or imprecision in the design itself (Vadde et al., 1994). That is, for probability theory to be applicable the assumption must hold that the possible performances (the 'events' - see section 2.3.1) are known and precisely defined (Zimmermann, 1983). Pomerol (1997) claims that in reality the possible performances may not actually be predictable with sufficient precision. Imprecision, however, cannot be modelled with probabilities (Efstathiou, 1984).

MAUT further assumes that well-specified, objective probability distributions can be established over all possible performances of each alternative with respect to each criterion (Keeney & Raiffa, 1976). Again, it does not seem to be absolutely clear whether this assumption is valid for practical decision-making in conceptual engineering design. Nojiri (1982) argues that one of the problems in conceptual design is lack of information, which may result in the established probability distributions actually being based on little more than best guesses. Ida et al. (1992) claim that it is not at all possible to identify any precise distribution. However, Lacksonen (1995) thinks that probability distributions can at least be estimated to a sufficient extent. Yet, Pomerol (1997) believes that even estimations seem not to be meaningful in many practical situations as probabilities are simply unknown and it remains unclear upon what information to base an estimation. According to Pomerol (1997) in many cases the decision-makers have subjective beliefs regarding the likelihood of some alternative's performance, but not an objective probability distribution, which would suggest that probability theory is not the appropriate modelling method for such cases.

Lotteries

MAUT uses the lottery technique to explicitly model the decision-maker's attitude towards uncertainty and to include the effect of uncertainty in the utility of an alternative. The ability to do so is seen as MAUT's particular strength (Weiss & Hari, 1997). Yet, it has been argued that the lottery technique is a non-intuitive approach, which causes difficulties for decision-makers (Barzilai, 1997) because determining the certainty equivalent for a lottery requires standards of accuracy that are seen as unrealistic and misleading (Efstathiou, 1984).

Summary

MAUT is an axiomatic decision-making method that is well supported by theoretical statistical foundations (Ivezic & Garrett, 1994). The method's particular strength is stated to be its ability to explicitly model the decision-maker's attitude towards uncertainty regarding performances and to include the effect of uncertainty in the utility of an alternative (Weiss & Hari, 1997).

A major point of criticism is that objective, probabilistic uncertainty, as modelled by MAUT, may not actually be the prevailing type of uncertainty occurring at decision-making situations in conceptual engineering design. It has also been criticised that MAUT has unrealistic demands for accuracy (Ivezic & Garrett, 1994).

2.3.4 Bayesian inferencing

As discussed in the previous section, decision-making under classical probability theory rests on the postulate that there is a known probability value for an event to happen which itself is the result of a randomising process. Yet, due to incomplete knowledge it may not be always possible to objectively assign these probability values. What can be done in such cases is to establish a subjective hypothesis and then update it constantly in the light of new information received. This is the underlying notion of Bayesian inferencing.

Bayesian inferencing is based on the concept of 'inverse probability' (Graham & Jones, 1988), which is the probability P of an event E_j having a given cause H_i (hypothesis), i.e. $P(H_i \mid E_j)$, rather than the probability P of a cause giving rise to an event E_j, i.e. $P(E_j)$. These two concepts are graphically shown in figure 2.12.

What is actually being measured with this inverse probability concept is the subjective change in a state of mind, which is the change in belief that a hypothesis is in fact true, rather than any change in an objective state (Graham & Jones, 1988).

To begin with, some hypotheses are made and prior probabilities are assigned to them. According to Bayes' 'Principle of Insufficient Reason' (Graham & Jones, 1988) these

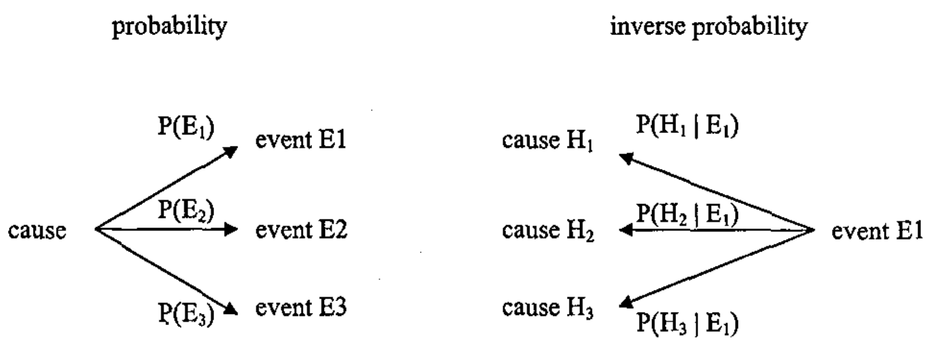


Figure 2.12: Probability and inverse probability.

prior probabilities are subjectively assigned and may be equal for all hypotheses if there is no 'Sufficient Reason' to the contrary.

Subsequently, tests are performed and on the basis of their outcome the prior probabilities are revised. The revised probabilities, called posterior probabilities, can be used as prior probabilities for further tests.

2.3.5 Application of Bayesian inferencing – how it works

A Bayesian approach towards modelling uncertainty is used by The Engineering Decision Support System (Herling, 1997). This system is based on a decision model whose structural elements are four pieces of information: alternatives, criteria, knowledge, and confidence. Alternatives are the proposed solutions to an issue. Criteria are the requirements, specifications, or constraints that measure the alternatives. Knowledge is an individual's personal experience about a specific alternative with respect to a particular criterion. Confidence is the level of surety about a propositional assertion that a specific alternative will satisfy a particular criterion. Using the Engineering Decision Support System (EDSS) requires the decision-maker to evaluate alternative-criterion pairs by stating a level of knowledge regarding the pair in question and a level of confidence that the alternative will satisfy the criterion. Apart from that, the decision-maker may assign importance weights for the criteria.

According to Bayes' 'Principle of Insufficient Reason' EDSS generally assigns a prior probability of 1.0 to any hypothesis 'criterion j is satisfied by alternative i'. Through gathering probabilities for knowledge and confidence from different decision-makers this prior probability may be updated (Herling et al., 1995). This is graphically shown in figure 2.13.

The decision-makers assign numeric probabilities to their knowledge and confidence levels indirectly by choosing a word from a knowledge and confidence word list. Each word in these lists has a numeric probability equivalent. The knowledge word list consists of the words and associated probability: 'experienced' (0.91), 'informed' (0.84), 'amateur' (0.78), 'weak' (0.66), and 'unknowledgeable' (0.57). The confidence

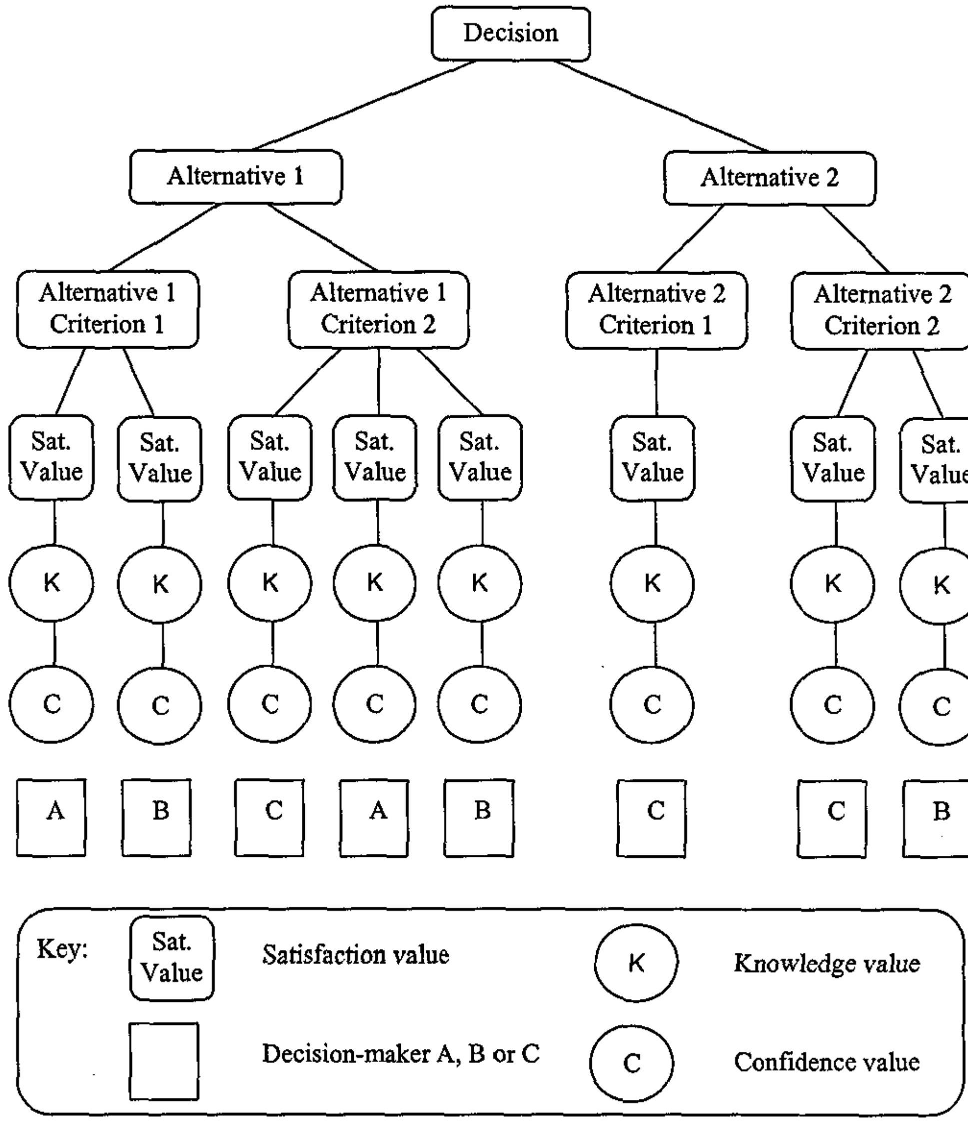


Figure 2.13: Nodal diagram of EDSS decision model (Herling, 1997).

word list consists of the words and associated probability values: 'perfect' (0.97), 'likely' (0.73), 'potential' (0.62), 'questionable' (0.42) and 'unlikely' (0.28).

EDSS calculates an 'overall satisfaction value' for an alternative with respect to all criteria (comprehensive evaluation) through a specifically developed weighted sum (Herling et al., 1995):

$$ParticipantSat(Alt_i) = \sum\nolimits_{criteria} V(NCri)\ SatValue_{each\ criterion}$$

$$SatValue_{each\ criterion} = \frac{\Pi_{Participant}[CK + (1-C)(1-K)]}{\Pi_{Particant}[CK + (1-C)(1-K)] + \Pi_{Participant}[C(1-K) + (1-C)K]}$$

where 'ParticipantSat(Alt_i)' is a particular participant's, i.e. decision-maker's, overall satisfaction value for alternative i, V(NCri) is this decision-maker's normalised criterion importance value, C is the confidence value, and K is the knowledge value.

It is important to note that, for the satisfaction value of each criterion, the probability of satisfaction increases as the knowledge and confidence increase and it decreases as knowledge increases and confidence decreases (Ullman et al., 1997).

As can be seen in the equation above, EDSS calculates the 'overall satisfaction value' of alternatives by aggregating all involved decision-makers' numeric probabilities for knowledge and confidence. However, only the criteria importance weightings of one individual decision-maker are used at a time. This means that if there are three decision-makers with different criteria weights, EDSS would calculate three 'overall satisfaction values' for each alternative reflecting each decision-maker's priorities regarding criteria.

2.3.6 Application of Bayesian inferencing – characteristics

Assumptions

EDSS is built upon a propositional, contrasted with statistical (randomness), probability framework and an absolute, contrasted with relative, evaluation technique (Herling et al., 1995).

The propositional probability framework rests upon a boolean position. This impacts the manner in which the alternatives and criteria are interpreted. The interpretation is either that the criterion is satisfied by the alternative, giving a probability of 1, or it is not satisfied by the alternative, giving a probability of 0. This means that the probabilities are about the certainty that the alternative will satisfy the criterion rather than about the degree to which satisfaction is achieved (Ullman et al., 1997). It is thus assumed that the decision-maker has a clear picture about what precisely it means that a particular criterion is satisfied, it is not sufficient to be merely clear about the direction of preference as when using pairwise comparisons. It is also assumed that it is not necessary to further differentiate the alternatives' performances as in 'both alternatives satisfy the criterion, but one is still preferred over the other'. EDSS's boolean position may be seen as unnatural: Yen and Tiao (1997) claim that criteria are naturally 'elastic', which means that their satisfaction is usually represented as a matter of degree, not in a black and white sense. EDSS further assumes that decision-makers are able to directly assign precise criteria importance weights.

Belief

Ullman and D'Ambrosio (1995) urge for the development of decision-making methods that model belief and its evolution. The EDSS is such a method. There are two aspects involved in modelling a decision-maker's belief: knowledge and confidence (Ullman et al., 1997). Knowledge is seen as a measure of the information held by a decision-maker about the attributes of the alternatives compared to the criteria. During design activities knowledge is generally increased (i.e. evolved) by building prototypes, performing simulations or finding additional sources of information (Ullman et al., 1997). Obviously, each of these activities to increase knowledge requires time and the commitment of resources.

Confidence quantifies the level of surety about a propositional assertion made that an alternative satisfies a criterion. To use a betting analogy, confidence defines how much of a bet to place on the alternatives success in meeting the criterion (Herling et al., 1995). Better knowledge will increase the confidence in some alternatives and reduce it in others (Ullman et al., 1997). Although the EDSS requires numeric probabilistic input for knowledge and confidence it requires little probability estimation for the decision-makers themselves. The decision-makers only need to choose words from a standard list to express their knowledge and confidence levels.

There may be conflict in viewpoint, knowledge and confidence among team members. The EDSS does not force all involved decision-makers to agree upon a unified belief. Instead, it supports conflicting beliefs and is based on each decision-maker's individual knowledge and confidence levels (Herling et al., 1995).

As the EDSS does not involve pairwise comparisons, but is based on direct, absolute evaluations it seems straightforward to include new alternatives and criteria. Thus, the EDSS supports a dynamic alternatives-space and a dynamic criteria-space.

Summary

The EDSS is capable of modelling subjective belief regarding the performance of alternatives with respect to criteria. Belief is expressed through assigning probabilities to the decision-maker's knowledge and confidence. It is not required to unify belief across teams, instead the method supports conflicting beliefs. The calculations are based on each decision-maker's individual knowledge and confidence levels (Herling et al., 1995). As the method applies direct, absolute evaluations it seems particularly straightforward to include new alternatives and criteria to the existing set.

A criticism addresses EDSS's incapability to differentiate between levels of performance with respect to criteria beyond the boolean 'satisfaction/no-satisfaction' statements. Apart from this, it requires the decision-maker to be clear about what precisely it means that a particular criterion is satisfied and it requires the decision-maker to directly assign precise criteria importance weights.

2.3.7 Dempster-Shafer theory

Some people regard Bayesian approaches for modelling belief as inappropriate, because these approaches treat belief like probabilities (Tanimoto, 1990). A strictly probabilistic representation of the belief that a hypothesis H is true, i.e. P(H), implies that the belief of this hypothesis being false is 1 - P(H), i.e. P(H) + P(~H) = 1. However, this implication is invalid in many practical cases. For example: There could be the hypothesis 'Liebfraumilch (a type of German wine) causes headache'. It is unlikely that somebody who does not even know what Liebfraumilch is would believe in this hypothesis. Yet, it is fair to say that this person would not believe in the opposite, either. Denoting this person's degree of belief by P(H), it is reasonable to assign both P(H) and P(~H) a value of 0. So, P(H) + P(~H) < 1.

The Dempster-Shafer theory allows for modelling belief without assuming that P(H) + P(~H) = 1. This theory distinguishes between ignorance, as in the example above, and uncertainty in the form of probability assignments for and against a hypothesis.

Probability assignments for and against a hypothesis can be represented by belief and disbelief, respectively. According to the Dempster-Shafer theory this can be expressed in form of a belief interval whose lower limit is called the support of the hypothesis H, i.e. belief B(H), and whose upper limit is called the plausibility of the hypothesis, PL(H). The plausibility of a hypothesis H is defined as unity minus the belief in the hypothesis' complement, i.e.:

$$PL(H) = 1 - B(\sim H).$$

Plausibility is a measure of whether the hypothesis could be true and must always be greater than or equal to the belief in the hypothesis. Some particular examples of belief intervals [B(H), PL(H)] and their meanings are introduced in table 2.5.

Interval	Meaning
[0, 1]	No belief in support of the hypothesis and no belief against it, i.e. total ignorance
[0, 0]	No belief in support of the hypothesis and maximal belief against it, i.e. hypothesis is believed to be false with certainty
[1, 1]	Maximal belief in the hypothesis and minimal belief against it, i.e. the hypothesis is believed to be true with certainty
[0.2, 1]	Partial belief in the hypothesis and minimal belief against it;
[0, 0.8]	Minimal belief in the hypothesis and partial belief against it;
[0.2, 0.8]	Partial belief in the hypothesis and partial belief against it, i.e. there is some evidence for and some evidence against the hypothesis

Table 2.5: Examples of belief intervals and their meaning.

The belief interval [B(H), PL(H)] is also referred to as the confidence in H, while the quantity PL(H) - B(H) is referred to as the uncertainty in H (Patterson, 1990). This may actually be regarded as 'an uncertainty about an uncertainty'. For example, B(H) = PL(H) = 0.7 implies that there is no uncertainty about the hypothesis' uncertainty of 0.7 since PL(H) - B(H) = 0. The belief interval may be updated through gathering evidence for or against H.

2.3.8 Application of Dempster-Shafer theory – how it works

The Dempster-Shafer theory is capable of modelling the refinement of beliefs with the accumulation of evidence. That is, the (subjective) probability of a hypothesis will be improved if more pieces of evidence can be found that support the hypothesis (Sen &

Yang, 1995). A possible hypothesis may be that the performance of an alternative is considered 'good' with respect to a particular criterion. If this criterion can be decomposed into a number of sub-criteria, then statements about the alternative's performance with respect to these sub-criteria are evidence for the criterion in the higher level.

A method for multi-criteria decision-making using Dempster-Shafer's theory has been introduced by Yang and Sen (1997). Their method supports evaluating alternatives by using hierarchically decomposed criteria. Subjective, qualitative, judgements with complete or incomplete uncertainty can be taken into account.

A criteria hierarchy including importance weights (in brackets) is given as an example in figure 2.14. The design problem referred to in figure 2.14 is the retro-fit of a stabilising structure for a sea going RoRo-Ferry. A structure is sought which changes the ferry's characteristics as little as possible. This example is taken from Yang and Sen (1997).

As delineated in figure 2.14, a criterion, such as 'daily operations requirements', can be associated with several lower-level criteria, which can be directly evaluated. Yang and Sen use linguistic statements such as 'good', 'very good' or 'poor' for evaluating

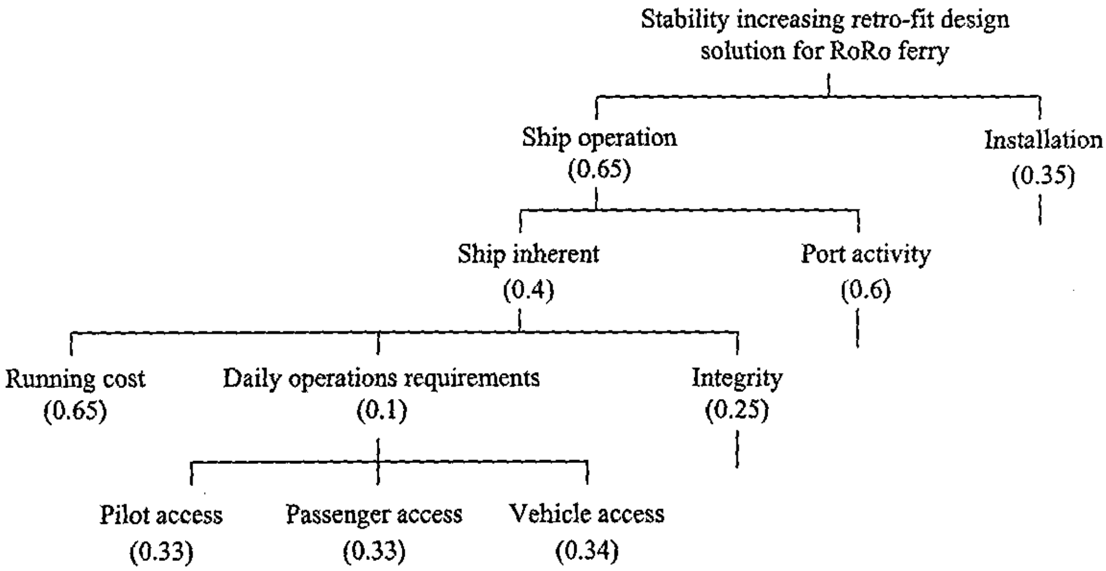


Figure 2.14: Hierarchical criteria structure.

the alternatives' performances. This is seen as a particularly straightforward evaluation approach especially for qualitative, subjective performances. The decision-maker may not even be absolutely sure about such a linguistic statement and may want to express a particular level of confidence in it. The challenge is then how to synthesise the statements and the associated confidence level regarding the lower level criteria to obtain a statement for the upper level criteria, including an indication of the decision maker's confidence.

Yang and Sen's method provides a way for synthesising 'confidence-enriched' evaluations by means of evidence combination for multiple criteria. The combined evidence supports a particular hypothesis. With respect to the example in figure 2.14, possible hypotheses may be that the performance of a specific alternative with respect to the criterion 'daily operations requirements' can be evaluated as e.g. 'major' or 'minor'. This means, the ferry's characteristics would be changed to a 'major' or 'minor' extent if the alternative in question was implemented. Evidence for these hypotheses is given by the evaluation of the alternative's performance with respect to the lower-level criteria.

In Yang and Sen's model, which refers to the example hierarchy given above, a set H of evaluation grades is defined which serves as a universe of discourse for the decision-maker. Within this set H the grade H_{n+1} is preferred to the grade H_n:

$$\begin{aligned} H &= \{H_1 \quad H_2 \quad H_3 \quad H_4 \quad H_5\} \\ &= \{Fundamental \quad Major \quad Moderate \quad Minor \quad None\} \end{aligned}$$

This set can be quantified by:

$$\begin{aligned} p\{H\} &= (p(H_1) \quad p(H_2) \quad p(H_3) \quad p(H_4) \quad p(H_5)) \\ &= (1 \quad 0.7 \quad 0.5 \quad 0.3 \quad 0) \end{aligned}$$

When evaluating an alternative's performance with respect to a low-level criterion, a decision-maker selects one or a number of evaluation grades and associates each selection with a confidence value $\beta_{kj}^{n}(A_r)$, where k is the high-level criterion, j is the low-level criterion, n is the evaluation grade supported, and A_r is the alternative in question. An evaluation with associated confidence is a piece of evidence $e_k^j(A_r)$. Each of these evidences implies a probability, which indicates the extent to which the evidence supports the hypothesis that a particular evaluation grade applies to the high-level criterion. This probability m_{kj}^{n} depends on the confidence that the decision-maker assigns to a particular evaluation grade when giving evidence, i.e. when addressing the low-level criteria, and on the importance of this evidence:

$$m_{kj}^{n} = \lambda_{k}^{j} \beta_{kj}^{n}(A_r),$$

where λ_{k}^{j} is the normalised relative importance of the evidence e_{k}^{j}.

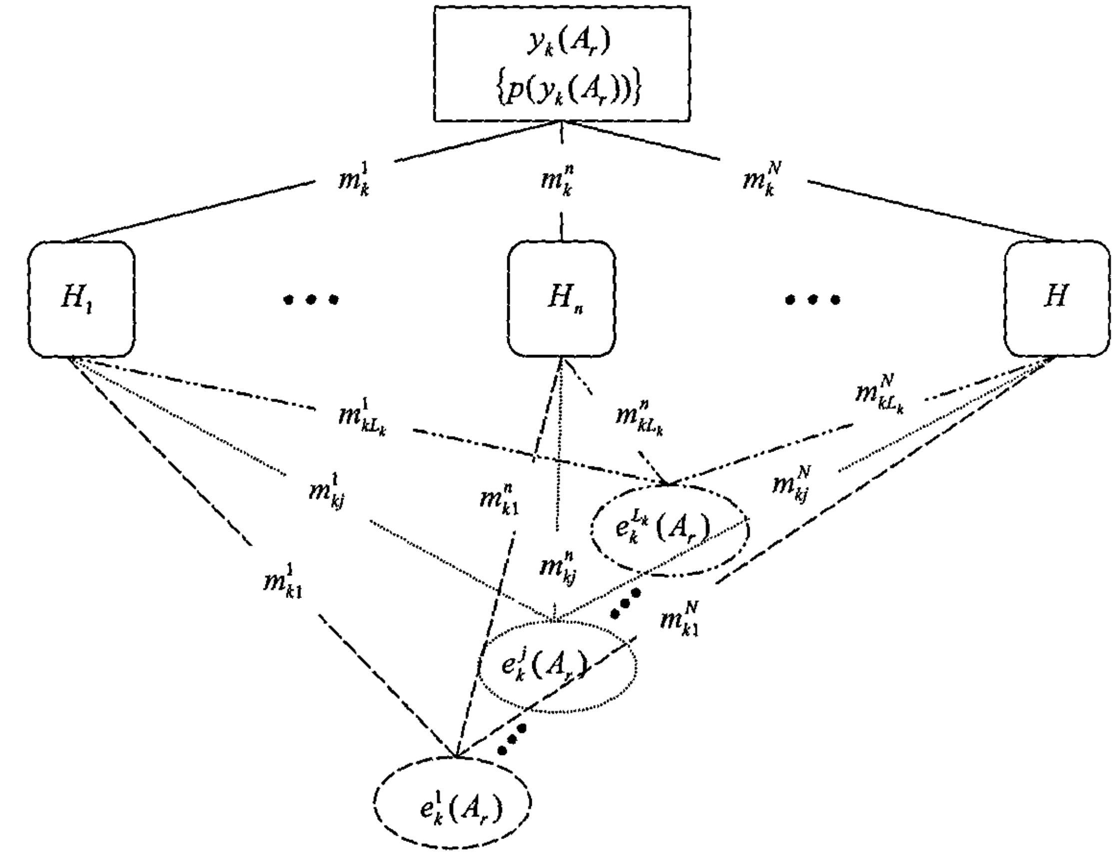
A decision-maker must assign a confidence level to a single evaluation grade or to a number of grades. The sum of all the confidence levels with respect to one criterion must not exceed unity, i.e.:

$$\sum_{n=1}^{N} \beta_{kj}^{n}(A_r) \leq 1,$$

where n=1, ..., N covers the entire set of evaluation grades. If a decision-maker assigns for example a confidence of 0.5 to the grade 'Moderate' and a confidence of 0.3 to the grade 'Minor', the confidence levels would not sum up to unity, i.e. $0.5 + 0.3 < 1$. Thus, the decision-maker's uncertainty is not complete. This means that there is some remaining belief unassigned after commitment of belief to all individual evaluation grades. The degree of this incomplete uncertainty is not assigned to individual evaluation grades, but to the whole set of evaluation grades H. The evidential reasoning approach works in a way that the degree of incomplete uncertainty will be reduced with the accumulation of evidence (Yang & Sen, 1997).

The overall probability assignment m_{k}^{n} to which the performance of an alternative A_r with respect to a high-level criterion y_k is confirmed to be of a grade H_n by the whole set of low-level criteria is obtained by combining the whole set of basic probability assignments m_{kj}^{n} ($j = 1, \ldots, L_k$). This combination and the combination of all incomplete uncertainties is carried out using the evidential reasoning algorithm which, due to its high level of complexity, will not be discussed here. Details about the mathematics of this methodology can be found in Yang and Singh (1994) and Yang and Sen (1994, 1997). A graphical model of Yang and Sen's evidential reasoning approach for a number of low-level criteria supplying evidence for one high-level criterion is shown in Figure 2.15. The same way of hierarchically propagating evidence as depicted in figure 2.15 can be applied through the entire criteria hierarchy. Then, the criterion y_k becomes a piece of evidence supporting hypotheses on the criteria one level above.

Using this approach it is also possible to model performances which are evaluated in numerical form. Such evaluation can be transformed through normalisation so that it will lie within the closed interval [0, 1]. Then it can be modelled in the same way as a subjective judgement, on the given universe of discourse, with complete uncertainty. Suppose y_{ij} is the evaluation of alternative i on criterion j with $0 \leq y_{ij} \leq 1$. If



Key:

y = criterion
A_r = alternative in question.
H_n = evaluation grade
k = high-level criterion
j = low-level criterion

$e_k^j(A_r)$ = evaluation of low-level criterion (piece of evidence)
m_{kj}^n = probability indicating extent to which $e_k^j(A_r)$ supports H_n
m_k^n = combined probability indicating extent to which all evidence supports H_n for y_k

Figure 2.15: A model of Yang and Sen's (1997) evidential reasoning approach.

$p(H_{n+1}) \le y_{ij} \le p(H_n)$ then it can be stated that alternative i on criterion j is evaluated to H_n and H_{n+1} with the confidence degrees of $(1 - \beta)$ and β, respectively, where β is given by:

$$\beta = \frac{p(H_n) - y_{ij}}{p(H_n) - p(H_{n-1})}.$$

To convey some clearer understanding about the type of inputs to and outputs from this method, tables 2.6 and 2.7 give an example with respect to the hierarchical attribute structure as shown in figure 2.14. In table 2.7 it is shown that the output of the evidential reasoning approach can be quantified as well. Details on how to calculate this value can be found in Yang and Sen (1997). Because an alternative is sought that changes the ferry's characteristic as little as possible a small value is preferred over a higher one.

Alternative	Evidence for the state of attribute y_1 (Daily operations requirements) $e_1^j(A_r)$	Subjectively assigned confidence values $\beta_{1j}^n(A_r)$ for the various evaluation grades H_n						unassigned uncertainty H
		Fundamental H_1	Major H_2	Moderate H_3	Minor H_4	None H_5	Weight	
Alternative A_1	Pilot access $e_1^1(A_1)$					1	0.33	0
	Passenger access $e_1^2(A_1)$			0.1	0.4	0.45	0.33	0.05
	Vehicle access $e_1^3(A_1)$			0.15	0.7	0.1	0.34	0.05

Table 2.6: Example inputs for Yang and Sen's method.

Alternative	Through evidential reasoning algorithm calculated probabilities m_1^n for the various evaluation grades H_n for attribute y_1 (Daily operations requirements)					Unassigned uncertainty H	Quantification
	Fundamental H_1	Major H_2	Moderate H_3	Minor H_4	None H_5		
Alternative A_1			0.0446	0.2970	0.6112	0.0472	0.1350

Table 2.7: Example outputs from Yang and Sen's method.

2.3.9 Application of Dempster-Shafer theory – characteristics

Assumptions

Yang and Sen's method requires the decision-maker to directly assess the absolute performance, in contrast to comparisons, of the alternatives with respect to criteria. This implies the assumption that a decision-maker 'knows' about absolute effectiveness rather than mere directions of preference. Yet, it is not assumed that an alternative's performance can be evaluated upon a quantitative scale. Instead, the decision-maker has at his/her disposal a predefined linguistic universe of discourse for qualitatively expressing performance evaluations. However, the linguistic expressions are transformed into precise probabilities, i.e. crisp numbers. It was already discussed in section 2.2.4 that the meaning of such transformations may be questioned. The decision-maker is also required to express his/her confidence in the performance evaluations by assigning precise probabilities.

A decision-maker using Yang and Sen's method needs to assign criteria importance weights. A linguistic universe of discourse may be used to assign these weights. The weights are not elicited by comparison, but through direct assignment. The universe of discourse limits the decision-maker to the predefined linguistic terms. These terms are then transformed into precise numeric quantities as for the performance evaluations. One may wonder whether these quantities actually express what the decision-maker had in mind when he/she assigned a linguistic term. This means that the numeric terms may pretend a level of precision that is not implied by choosing a particular linguistic expression.

Belief

The Dempster-Shafer theory allows for modelling belief and disbelief. In Yang and Sens's method this does not seem to be fully implemented as it is impossible to explicitly express 'disbelief' or 'plausibility'. Yet, it is possible to model unassigned belief, which means that lack of confidence or ignorance can be expressed.

Yang and Sen's belief model is based on the accumulation of evidence through a criteria hierarchy (Yang & Sen, 1997). Many researchers consider the support for establishing criteria hierarchies as a particularly advantageous feature for decision-making methods. This was already discussed in section 2.2.4.

Summary

Yang and Sen's method allows for the representation of subjective uncertainty. This uncertainty is modelled as belief, represented by a level of confidence in a qualitatively described performance evaluation. The level of confidence needs to be expressed in precise numeric probabilities. There is also an indication about the completeness of the uncertainty. This can be interpreted as a decision-maker's uncertainty about their confidence in the performance evaluations.

For the evaluation of performances and for the assignment of criteria importance weights a standard linguistic universe of discourse may be used. This simplifies evaluations. Yet, the transformation of linguistic statements into precise quantities for numeric processing may be questionable.

2.3.10 Fuzzy set theory

Fuzzy set theory was introduced by Zadeh (1965) as a framework which provides a natural way of dealing with problems in which the source of uncertainty is the absence of sharply defined criteria of class membership, i.e. vagueness or imprecision. Zadeh (1965) defines a fuzzy set as a class of objects with a continuum of grades of membership. Examples of such classes are the class of *fast* cars, the class of *tall* men, or the class of numbers which are '*about* 2'.

A fuzzy set is characterised by its membership function, which determines the degree of membership for each element of the universe on which the fuzzy set is defined. The degree of membership is expressed by associating each element of the universe with a real number in the interval [0, 1]. If the membership function associates a particular element with the membership value 0 then this element is not a member of the fuzzy set. The membership value 1 expresses full membership and any value between 0 and 1 expresses partial membership. For example: Let the universe X with the generic element x be the real line R and let A be a fuzzy set of numbers which are 'about 2'. The linguistic concept 'about' introduces subjective vagueness, which can be modelled by the fuzzy set A through its membership function $\mu_A(x)$. Some representative values of this membership function may be $\mu_A(1.5) = 0$, $\mu_A(1.7) = 0.4$, $\mu_A(1.9) = 0.8$, $\mu_A(2) = 1$, $\mu_A(2.1) = 0.8$, $\mu_A(2.3) = 0.4$, $\mu_A(2.5) = 0$. The continuous form of the membership function, covering all elements of R rather than a few representatives only can be expressed as follows:

$$\mu_A(x) = \begin{cases} 0 & \text{if } x \leq 1.5 \\ 2x - 3 & \text{if } 1.5 < x < 2 \\ 1 & \text{if } x = 2 \\ -2x + 5 & \text{if } 2 < x < 2.5 \\ 0 & \text{if } x \geq 2.5 \end{cases}$$

Fuzzy set A representing the subjective concept 'about 2' is graphically shown in figure 2.16.

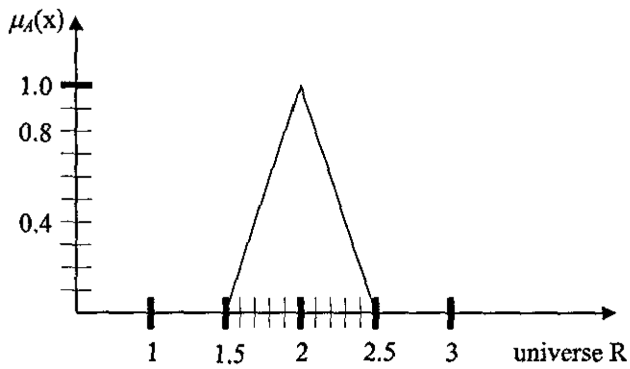


Figure 2.16: Graphical representation of fuzzy set A, 'about 2'.

In classical set theory a particular element can only be either a full member of a set or not a member at all. This implies that the vague term 'about' cannot be modelled by classical sets, which are also called crisp sets (Zimmermann, 1983). Hence, a particular strength of fuzzy sets lies in their ability to model vague (linguistic) terms.

2.3.11 Application of fuzzy set theory – how it works

The first application of fuzzy sets to decision-making was introduced by Bellmann and Zadeh (1970). They define a decision as the fuzzy set of alternatives resulting from the intersection of the goals and constraints. In their approach goals and constraints are fuzzy sets on the same universe, which is the alternative space. As an example one can consider the real line R as the universe X, the alternative space, with the generic element x. This means that the decision alternatives are represented by the real numbers. A fuzzy constraint C may state that x should not be much greater than 15 and a fuzzy goal G may state that x should be substantially larger than 10. The fuzzy sets modelling constraint C and goal G are, subjectively, characterised by the following membership functions:

$$\mu_C(x) = \begin{cases} 0 & \text{if } x \le 15 \\ -0.5x + 8.5 & \text{if } 15 < x < 17 \\ 0 & \text{if } x \ge 17 \end{cases} \qquad \mu_G(x) = \begin{cases} 0 & \text{if } x \le 10 \\ 0.1x - 1 & \text{if } 10 < x < 20 \\ 0 & \text{if } x \ge 20 \end{cases}$$

Constraint C as well as goal G shall be satisfied simultaneously which means that they are linked by the connective 'and'. 'And' is implemented in set theory by an intersection of sets. In fuzzy set theory an element's membership value in the set representing the intersection of two sets is the minimum of this element's membership value in either of the intersecting sets. For example, if the element 'a' has a membership value of 0.7 in the one set and 0.5 in the other set, its membership value in the intersection of these two sets would be 0.5. The intersection of the goal and the constraint above can be expressed as:

$$\mu_{G \cap C}(x) = \begin{cases} 0 & \text{if } x \le 10 \\ \min(-0.5x + 8.5,\ 0.1x - 1) & \text{if } 10 < x < 17 \\ 0 & \text{if } x \ge 17 \end{cases}$$

The fuzzy set, representing the intersection is characterised by $\mu_{G \cap C}(x)$. The best decision can now be found by selecting the alternative (x_i) with the maximal membership function in the intersection. The relationship between decision, goal, and constraint is graphically shown in figure 2.17.

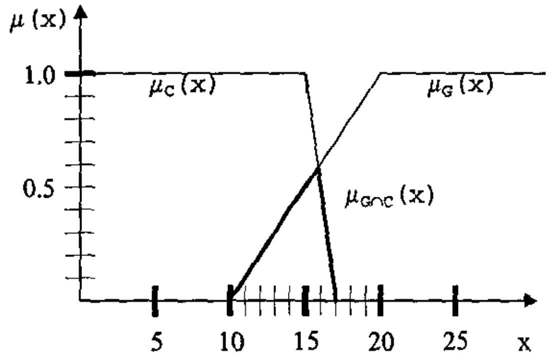


Figure 2.17: Relationship between fuzzy goal, fuzzy constraint, and decision.

It is questionable whether the particular approach as described above can be applied to engineering design decision-making. This is mainly because although vagueness which is inherent in the constraints and goals can be expressed, it is not possible to represent fuzzy information referring to the 'system state' (Jain, 1976), i.e. the performance of alternatives. With reference to the example above this means that it must be known, that one alternative is for example exactly 15 rather than about 15 in order to be able to determine a membership value in the intersection and thus a decision. However, since this first application was introduced, considerable research in fuzzy decision-making has been undertaken and a number of very different, partially rather complex, approaches have been introduced.

Thurston and Carnahan (1992) discuss the application of fuzzy set theory for engineering design decision-making. They apply fuzzy sets to represent linguistic variables. Using their method, a decision-maker may express him/herself linguistically with regard to the weights of criteria and the performance of alternatives with respect to these criteria. The linguistic variables, i.e. weights and performances, can assume values from a standard universe of discourse. This universe of discourse contains a particular, pre-defined, number of fuzzy sets with a linguistic meaning. Seven of these sets are used: very low (VL), low (L), low to middle (ML), middle (M), middle to high (MH), high (H), and very high (VH). Figure 2.18 shows the universe of discourse.

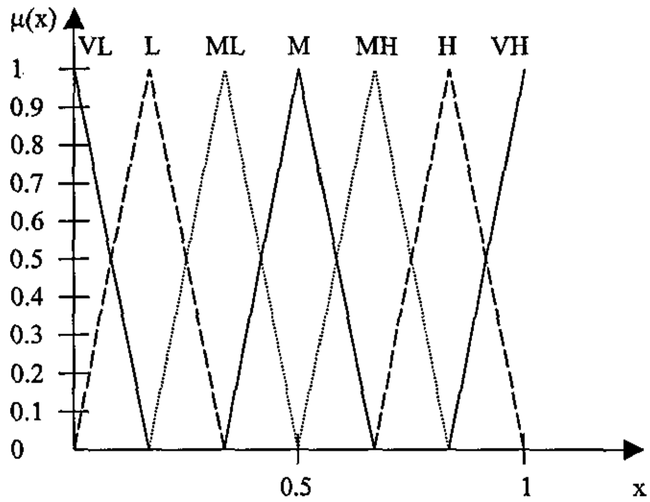


Figure 2.18: Membership functions for universe of discourse.

The fuzzy sets as shown in figure 2.18 quantify the vagueness of the linguistic statements on a scale between 0 and 1. By applying the mathematics of fuzzy set theory, in particular Zadeh's (1965) 'extension principle', it is possible to combine this kind of fuzzy information in a similar way as crisp information can be combined. This means that a decision-making aggregation function can be applied to establish a statement on the overall effectiveness of an alternative. As in the crisp case, criteria importance weightings and design solution performances are factors in this function. The result of aggregating fuzzy information will be fuzzy as well. A fuzzy overall rating of two alternatives A and B is shown graphically in figure 2.19.

A standard universe of discourse as in figure 2.18 may cause difficulties for the decision-maker if he/she would like to choose a fuzzy set representing a linguistic meaning in between two available, i.e. standard, sets offered by the universe. For example, the decision-maker may want to express a particular alternative's performance as being less than 'high', but more than 'medium high'. Carnahan et al. (1994) have introduced the 'fuzzy line segment' in order to overcome this difficulty.

By applying the fuzzy line segment a parameterised fuzzy set can be created which expresses a linguistic meaning in between two standard sets within the universe of discourse. The new set's position between the standard sets is determined by the parameter's value. This concept can be imagined as shown in figure 2.20. The parameter α determines a point on the line $\overline{ab}$ between two adjacent fuzzy sets.

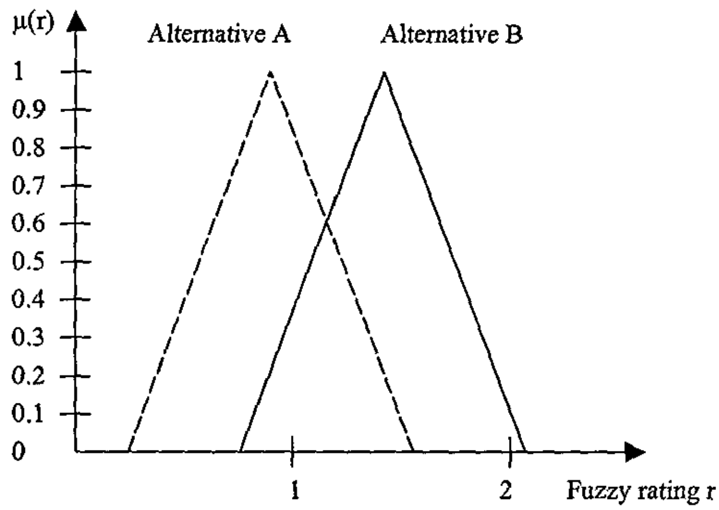


Figure 2.19: Fuzzy overall ratings of alternatives A and B.

Attached to this point is a fuzzy set that can be 'dragged' along this line by choosing values for α which determines the distance from a. The values for α may be in the interval [0, 1]. A value close to 0 lets the linguistic meaning of the new set be close to the one of its leftward neighbour whereas a value close to 1 lets the linguistic meaning of the new set be close to the one of its rightward neighbour. A value of 0.5 implies that the new set's meaning is exactly in the middle between the meanings of its two adjacent sets.

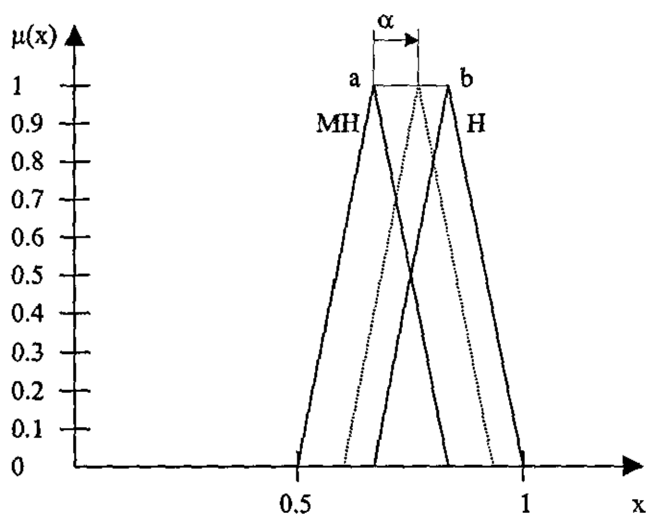


Figure 2.20: Fuzzy line segment between 'medium high' and 'high'.

The fuzzy line segment can also be applied for the fuzzification of performances that are already estimated in numeric rather than linguistic form (Carnahan et al., 1994). An example of such a numeric estimation is 'about 50 miles'. In such a case the universe of discourse is seen as a scale where the minimum possible numeric value is on x = 0 and the maximum possible numeric value is on x = 1. The fuzzy set representing the given value can then be placed at its exact position between the two sets, which represent the values just below and just above the given value.

Another way of fuzzifying numeric values is shown by Zimmermann and Sebastian (1995). As a feature of their decision-support system KONWERK they associate levels of noise intensity with linguistic meanings represented by fuzzy sets. It was suggested to apply membership functions that have a trapezoidal rather than a triangular shape, as shown in figure 2.21. This implies that there are intervals in which no distinction is made between different noise levels. Thus, it is easier to find a fuzzy set from the existing universe of discourse, which may represent a particular numerical value.

Also as a feature of KONWERK Müller and Sebastian (1997) have developed a way for adjusting the sensitivity level of a pre-defined, linguistic universe of discourse. KONWERK offers a range of linguistic universes of discourse, each one

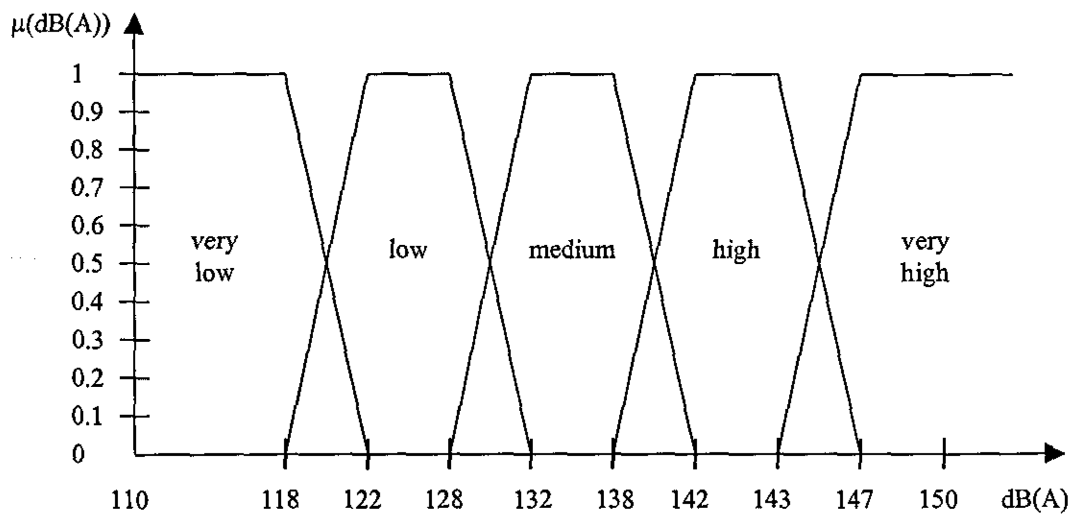


Figure 2.21: Trapezoidal membership functions for linguistic variables.

corresponding to a different sensitivity level. The lowest sensitivity level is represented by a universe that consists of only two sets: 'low' and 'high'. The second sensitivity level is generated by adding a set for the meaning 'medium', the third level by adding 'very low' and 'very high' and so on. The highest sensitivity level contains sets for 'extremely low', 'very low', 'low to very low', 'low', 'fairly low', 'more or less low', 'medium', 'more or less high', 'fairly high', 'high', 'high to very high', 'very high', and 'extremely high'. Figure 2.22 shows as example the levels two and three.

The appropriate level of sensitivity is chosen automatically by the decision support system. If a user expresses for instance the performance levels of all alternatives with respect to a particular criterion with the terms 'low', 'medium', or 'high', sensitivity level 2 would be selected. If, however, the term 'very high' or 'very low' appeared, level 2 would not be suitable any more and level 3 would be selected. As can be seen in figure 2.22, the membership function for a particular linguistic statement, such as 'medium' changes through the different levels of sensitivity. This implies that the actual meaning of a concept like 'medium' is interpreted in rather vague terms as well. The decision-maker should be aware of that.

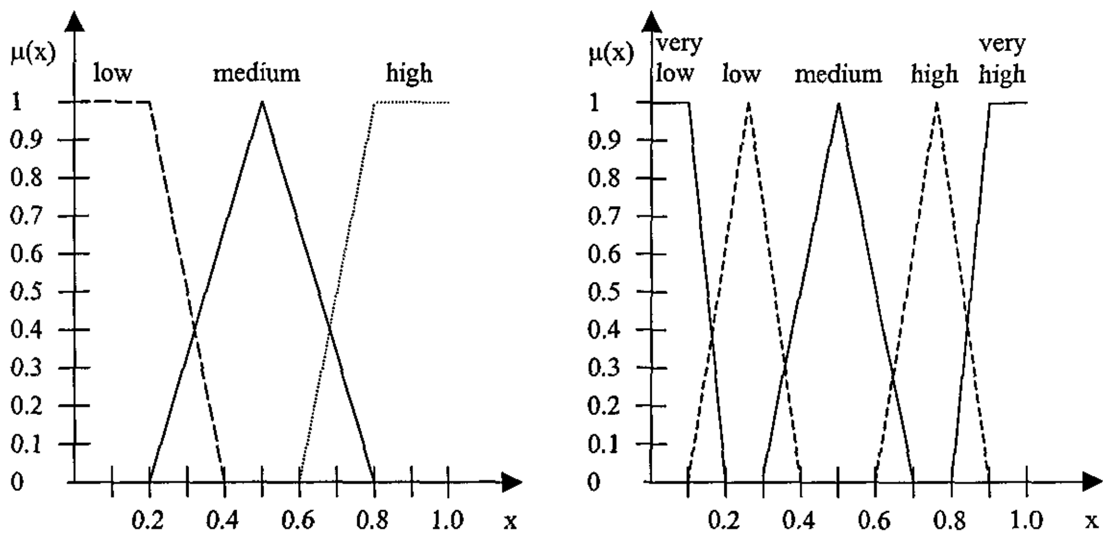


Figure 2.22: Universes of discourse with different sensitivities – level 2 (left) and level 3 (right).

Membership functions that differ from the triangular or trapezoidal shape, as described above, have been applied by Allen (1996). Her π-functions take on the shape of a bell and are defined by:

$$\mu_{D(x)} = \begin{cases} 0 & \text{if} \quad x \leq (p - \beta) \\ \left(\frac{2}{\beta^2}\right)(x - p + \beta)^2 & \text{if} \quad (p - \beta) < x < \left(p - \frac{\beta}{2}\right) \\ 1 - \left(\frac{2}{\beta^2}\right)(x - p)^2 & \text{if} \quad \left(p - \frac{\beta}{2}\right) \leq x \leq \left(p + \frac{\beta}{2}\right) \\ \left(\frac{2}{\beta^2}\right)(x - p - \beta)^2 & \text{if} \quad \left(p + \frac{\beta}{2}\right) < x < (p + \beta) \\ 0 & \text{if} \quad x \geq (p + \beta) \end{cases}$$

where p is the peak value, and β is the bandwidth. That is, 2β is defined as the distance between the two 'ends' of the membership function. An example of such a membership function is given in figure 2.23. Such functions may more accurately interpret people's understanding of linguistic expressions than triangular or trapezoidal functions.

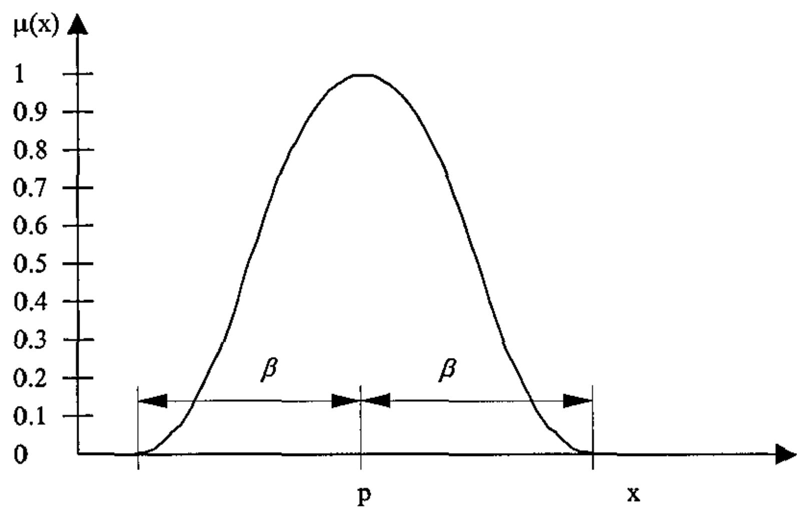


Figure 2.23: Π-membership function.

2.3.12 Application of fuzzy set theory – characteristics

Assumptions

Decision-making methods that model uncertainty by using fuzzy sets assume that the involved uncertainties are imprecision, ambiguity or vagueness (Vadde et al., 1994). According to Yao and Furuta (1986) these uncertainties may be distinguished from one another as follows:

- Imprecision is associated with incomplete information;
- Ambiguity is associated with using natural language, which is meaningful but not clearly defined.
- Vagueness is associated with inexact and/or ill-defined figures.

When a decision-maker is providing information, his or her uncertainty may be traced to internal as well as external sources. Internal uncertainty may be introduced through human reasoning processes (Efstathiou, 1984). External uncertainty may exist with respect to events that have not occurred yet, such as design concepts that have not been implemented.

With respect to internal uncertainty it is believed that decision-makers in many organisations do not always have a clear and precise understanding of their own goals and values (Pomerol, 1997). According to Zimmermann (1983) such human factors introduce imprecision, ambiguity or vagueness. This suggests that fuzzy methods are the appropriate modelling approach.

With respect to external uncertainty there has been a long-running debate concerning whether probabilistic methods or fuzzy methods are the appropriate modelling basis (Thurston & Carnahan, 1992). There are many researchers who advocate that actually there is no probability at all (Pomerol, 1997). Instead, decision-makers generally do not agree on or are not able to make sufficiently precise predictions. Law and Antonsson (1995) agree and argue that in conceptual design this is due to the intrinsic vagueness of an unfinished design description. Such unfinished design descriptions may only allow for predictions by subjective judgement rather than objective analysis (Conrath, 1973). It is claimed that human judgements are bound to have errors and biases (Pöyhönen & Hämäläinen, 2000) which then result in imprecision and vagueness (Zimmermann, 1983). Moreover, ambiguity may be introduced if human decision-makers express their predictions in linguistic terms rather than numerically. This again is the realm of fuzzy sets.

For a better understanding of how fuzzy uncertainty differs from probabilistic uncertainty one may interpreted fuzzy sets as possibility distributions in contrast to probability distributions (Ida et al., 1992; Zimmermann, 1983). The different meanings of a possibility distribution and a probability distribution may become clear in the following example taken from Zimmermann (1983): Let us consider the statement 'Paul eats X eggs for breakfast'. The possibility distribution po(X) and the probability distribution pr(X) might then be as shown in table 2.8.

X	1	2	3	4	5	6	7	8	9	10
po(X)	1	1	1	1	0.8	0.6	0.4	0.2	0	0
pr(X)	0.1	0.8	0.1	0	0	0	0	0	0	0

Table 2.8: Example of possibility and probability distribution.

Computing with words

Researchers (Vadde et al., 1994) have pointed out that during the initial stages of a design project, the conceptual design phase, usually very little 'hard' information is available. Law and Antonsson (1995) state that during this phase, the description of a design is completely vague or imprecise. Having only a vague description of a design implies that decision-makers are only able to make vague predictions regarding its expected performance. Moreover, Carnahan et al. (1994) state that usually the goals of the design project are only vaguely defined as well. Goals of the design project are related to meeting the given specifications and requirements. It is claimed that in 'real' design, designers must often ask questions to distinguish the underlying fuzzy constraint so that the final design will satisfy the customer's *actual* requirements even though it may violate the crisp constraint initially given (Antonsson & Otto, 1995). Finally, it is believed that designers can describe the importance of different goals, or decision-making criteria, only in vague, imprecise terms (Thurston & Carnahan, 1992).

In spite of all this vagueness and imprecision a decision must be made so that the subsequent detailing of promising alternatives can proceed. Developments in fuzzy set theory find direct application to this aspect of the engineering design process (Wood & Antonsson, 1987). Fuzzy sets permit modelling 'fuzzy' information as occurring during conceptual design in a way that it can be used for formalised analysis to facilitate decision-making support. A number of researchers believe, the fuzziness of performance predictions, the fuzziness of design goals and the fuzziness of criteria importance can all be represented in exactly the same way through the mathematics of fuzzy sets (Thurston & Carnahan, 1992; Antonsson & Otto, 1995).

'Fuzzy' information as occurring during conceptual design may often be communicated by linguistic expressions, such as 'high cost' or 'low weight', which are seen to reflect the designers' current best estimate of, for example, expected performances (Thurston & Carnahan, 1992). Ghotb and Warren (1995) think that the major advantage of fuzzy methods is that they can deal with linguistic input and thus express the 'exact feeling' of human decision-makers. Müller and Sebastian (1997) say that this ability makes them in many situations more convenient to use than methods requiring precise inputs. This should be the case in particular for aspects that naturally suggest linguistic assessments, such as 'comfort'. It is claimed that the use of semantics also facilitates reaching consensus among members of decision-making groups on issues such as the relative performance of alternatives (Thurston & Carnahan, 1992).

To formally process linguistic statements they are quantified through fuzzy sets. One may wonder whether a quantification of linguistic expressions can in fact convey linguistic meaning. Zimmermann (1983) argues that already by expressing their thoughts due to the 'poorness' of natural language people have to approximate their thought processes more or less well by selecting the verbal connective which is closest to their real connective. Thus, his argument is that the actual approximation takes place already when verbally expressing thoughts and not so much when turning these expressions into a mathematical model. Zimmermann (1996) considers 'computing with words' as the big challenge for fuzzy technology.

Quantitative input

Pure linguistic statements may be used to express qualitative factors. For expressing imprecise quantities, fuzzy numbers may be used (Ghotb & Warren, 1995). As an example, the fuzzy number 'about 2' was discussed in section 2.3.10.

Müller and Sebastian (1997) explain that whereas at the concept stage the description of a design is nearly completely vague or imprecise, this imprecision is reduced through the design process until ultimately the final description is precise. The process of reducing imprecision can be represented through fuzzy sets. Allen (1996) states that crisp, non-fuzzy, numbers can be interpreted as fuzzy numbers with a bandwidth of zero. She further states: Operations on crisp numbers can be performed on fuzzy numbers with a bandwidth of zero; this provides mathematical continuity between crisp and fuzzy numbers. In design decision-making situations that involve both crisp and fuzzy numbers, the entire problem is modelled with fuzzy numbers and fuzzy arithmetic is used. In successive redesign, as the design is developing, presumably, more and more is known about the object being designed and there is less and less uncertainty regarding its expected performance. Thus, it will be appropriate to use smaller and smaller bandwidths in the membership functions expressing expected performance, however, the mathematical formulation will remain the same.

Membership functions

A difficulty caused by fuzzy sets regards how to measure 'fuzziness' or the 'degree of membership' (Zimmermann, 1983), i.e. how to determine the exact membership function. Thurston and Carnahan (1992) suggest that for practical purposes it is often best to restrict the fuzzy sets to triangular membership functions. For calculations it is then possible to apply the approximate formulas introduced by Dubois and Prade (1978). For more complex functions the calculations involved become rather complicated. Yet, even triangular functions may have various 'shapes', in particular various bandwidths. Zimmermann (1983) believes that exact functions could be empirically determined. This, however, does not appear to be too straightforward. Zimmermann (1983) concludes that the flexibility of fuzzy sets seems to be one of its strengths and weaknesses at the same time.

Defuzzification

The application of fuzzy sets and fuzzy mathematics for the calculation of alternatives' overall effectiveness results in fuzzy outputs. Yet, what designers eventually require is a crisp decision rather than a fuzzy one. There are a number of analytical approaches for transforming fuzzy outputs into crisp rankings (Thurston & Carnahan, 1992). However, as Ghotb and Warren, (1995) argue, the difficulty is that there is no single approach which is generally considered superior.

Allen (1996) gives some guidelines about how to approach a practical ranking of fuzzy sets. A problematic case with two overlapping sets that need to be ranked is shown as an example in figure 2.24. These two sets represent the fuzzy representation of two alternative concepts' overall effectiveness. For such a case Allen (1996) suggests that a designer may choose either concept or may choose to do further engineering to narrow the bandwidths of the fuzzy result sets. She recommends that fuzzy sets are ordered first on the basis of their modal values and, if two fuzzy sets have the same modes, then fuzzy sets with narrow membership functions are ordered above those with a broad membership function. In any case, she argues, a designer has a graphical representation of the situation and has information on which to base his/her decision. The effects of uncertainty are immediately apparent to a designer who may then choose concepts for further development or may decide to reduce the uncertainty associated with particular concepts by seeking more precise information.

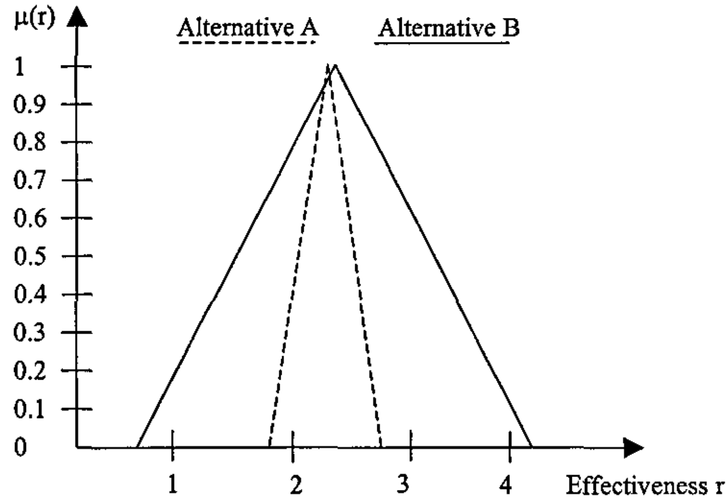


Figure 2.24: Fuzzy sets representing overall effectiveness.

Summary

Fuzzy set theory is a framework which provides a natural way of dealing with problems in which the source of uncertainty is the absence of sharply defined criteria of class membership (Zadeh, 1965).

Fuzzy sets permit the modelling of vague, imprecise and/or ambiguous information in a way that it can be used for formalised analysis to facilitate decision-making support. It is claimed that the fuzziness of performance predictions, the fuzziness of design goals and the fuzziness of criteria importance can all be represented in exactly the same way through the mathematics of fuzzy sets (Thurston & Carnahan, 1992; Antonsson & Otto, 1995). The ability of fuzzy methods to deal with linguistic input and thus express the 'exact feeling' of human decision-makers is considered a major advantage (Ghotb & Warren, 1995). Fuzzy sets can also be used for modelling imprecise and precise quantitative aspects of decision-making situations (Allen, 1996).

Difficulties caused by fuzzy sets regard (i) how to measure 'fuzziness' or the 'degree of membership' (Zimmermann, 1983), i.e. how to determine the exact membership function, and (ii) how to 'defuzzify' fuzzy results to obtain crisp rankings. Some practical guidelines of how to deal with these difficulties have been suggested by Thurston and Carnahan (1992), Zimmermann (1983) and Allen (1996).

The particular power of a decision that is based upon fuzzy analysis rather than crisp analysis can be summarised by a statement taken from Schwartz (1962, p357): "An argument which is only convincing if it is precise loses all its force if the assumptions on which it is based are slightly changed, while an argument, which is convincing but imprecise may well be stable under small perturbations of its underlying axioms".

2.4 Discussion

Within the previous sections a number of different methods have been introduced which can be applied to support decision-making situations. Section 2.2 concentrated on methods for multi-criteria decision-making and section 2.3 concentrated on methods for modelling uncertainty. Within the following sections the main characteristics of these methods will be summarised and related to each other. Finally a research gap will be identified.

2.4.1 Characteristics of methods for multi-criteria decision-making

Section 2.2 introduced three major approaches towards multi-criteria decision-making: multi-attribute value theory (MAVT) and direct scoring, the analytic hierarchy process (AHP), and the wider known outranking methods Electre, Promethee as well as decision-making with Quality Function Deployment (QFD) and Pugh's method of controlled convergence.

A major difference between these approaches is the way in which the alternatives' performances are evaluated with respect to criteria. MAVT evaluates the alternatives' performances by mapping them onto an independent scale, the individual effectiveness functions, whereas the AHP and the outranking methods evaluate the alternatives' performances relative to each other by comparison.

MAVT needs for the establishment of effectiveness functions the defined ranges of all alternatives' performances. Then the actual function is required that maps all possible performances within the defined range onto a numeric scale representing effectiveness. If direct scoring methods are used it is not necessary to be explicit about a precise effectiveness function; scores may be estimated on an interval scale. Yet, as these scores represent independent, absolute evaluations, a general idea about the possible range of performances and some form of implicit effectiveness function is still required. MAVT and direct scoring methods create not only a relative order of alternatives, but also an absolute measure of overall worth for each alternative.

If the alternatives' performances are evaluated relative to each other by comparison, as in the AHP and the outranking methods, no effectiveness function is needed. It is only necessary to know whether the performance of one alternative is preferred over the performance of another alternative and to what extent. As comparisons are of a relative nature, the final result will only be a relative order of alternatives, a ranking rather than a rating, and there will be no indication on whether even the highest ranked alternative is very effective overall. This means that, in contrast to MAVT, comparison methods do not generate an absolute measure of overall worth.

A main difference between the AHP and the outranking methods is the sensitivity of the relative performance evaluations. The AHP requires the decision-maker to exactly quantify by how much one alternative's performance is preferred over another alternative's performance. When using an outranking method a decision-maker is not required to quantify comparisons, but it is sufficient to express them in qualitative terms.

As a consequence of using quantified comparisons the AHP's overall rankings are quantitative as well, which means that differences between alternatives regarding overall worth are indicated. This is in contrast to rankings as established with outranking methods. They are of a qualitative, ordinal nature.

Another major difference between MAVT, AHP, and outranking methods is the way in which criteria weighting factors, representing the importance of criteria, are elicited. In MAVT these weights are usually directly assigned. In the AHP the weights are elicited through iterative pairwise comparisons. The outranking methods usually apply direct weight assignments as in MAVT. An exception is Pugh's method of controlled convergence, which implicitly assumes that all criteria are of equal importance and therefore does not require the elicitation of criteria weights.

2.4.2 Characteristics of methods for modelling uncertainty

Section 2.3 introduced four major approaches towards modelling uncertainty: probability theory, Bayesian inferencing, the Dempster-Shafer theory, and fuzzy set theory.

A major difference between these four approaches is the type of uncertainty being modelled. Probability theory models the objective likelihood regarding an expected occurrence of a particular event. Bayesian inferencing and the Dempster-Shafer theory model the subjective belief in the occurrence of a particular event. The difference between Bayesian inferencing and the Dempster-Shafer theory is that the former only models belief, whereas the latter models belief as well as disbelief. Fuzzy set theory, on the other hand, models imprecision, vagueness or ambiguity in the description of an event. This theory is not concerned with the likelihood of an event's occurrence.

In section 2.3 the characteristics of particular decision-making methods were discussed that are based on the above four approaches for modelling uncertainty. These were multi-attribute utility theory (MAUT), the Engineering Decision Support System (EDSS), Yang and Sen's method, and different fuzzy methods.

MAUT is based on probability theory. What is usually modelled with MAUT is uncertainty with respect to the alternatives' expected performances, but not uncertainty with respect to criteria importance weights. MAUT assumes that well-specified, objective probability distributions can be established over all possible performances of each alternative with respect to each criterion.

EDSS is based on Bayesian inferencing, modelling belief. It is built upon a framework that rests on a boolean position. This means that the decision-maker's belief, in EDSS called 'confidence', is about the certainty that the alternative will satisfy the criterion rather than about the degree to which satisfaction is achieved. The EDSS does not model uncertainty with respect to criteria importance weights.

Yang and Sen's method is based on Dempster-Shafer theory. Using this method it is possible to model belief, which is represented by a level of confidence in a qualitatively described performance evaluation. This performance evaluation does not have to be boolean as in EDSS. Yang and Sen's method does not seem to explicitly model 'disbelief' or 'plausibility' as in Dempster-Shafer theory. But, it is possible to model unassigned belief, which can be interpreted as a decision-maker's uncertainty about their confidence in the performance evaluations. In contrast to EDSS, when using Yang and Sen's method a decision-maker needs to express his/her confidence as precise, numeric probabilities. Yang and Sen's method does not model uncertainty with respect to criteria importance weights.

A number of methods were introduced that are based on fuzzy set theory. These methods model the fuzziness of performance predictions, the fuzziness of design goals and the fuzziness of criteria importance weights. A particularly distinctive characteristic of fuzzy methods is their ability to model the inherent ambiguity of natural language, which enables 'computing with words'. Fuzzy methods do not model the probability of performance predictions or the belief in evaluations.

2.4.3 Research gap

The general aim of the discussions in sections 2.2 and 2.3 was to identify available support methods and to develop an understanding of the field. In particular, I tried to generate insight into the 'manner of thinking and reasoning' i.e. the assumptions and the underlying principles upon which decision-making methods are based. I call these assumptions and principles the methods' characteristics.

Are all methods equally effective?

Through discussing the methods' characteristics and through relating them to each other, as in the summary sections 2.4.1 and 2.4.2, a number of aspects emerged that seem to play a role in structure-related methodological decision-making support for selection situations (see section 1.3.3, focus on selection-type decisions). It became clear that the methods are very different from each other and it could be realised with respect to what aspects they differ in which way. Knowing about this, one may now wonder whether all these methods are still equally effective for supporting decision-making in conceptual engineering design or whether some are more suitable than others. This question has, in fact, been asked before (Roozenburg & Eekels, 1995).

Indications about the usefulness of particular methods may be gained from information regarding the extent with which they are actually used. Yet, when I tried to gather such information I became aware of many researchers who claim that none of the available methods has gained wider use in industrial design practice (Ullman & D'Ambrosio, 1995; Žavbi & Duhovnik, 1996). Instead, designers apparently use their intuition, experience and instinct (Cziulik & Driscoll, 1997). Some researchers (Liberatore and Stylianou, 1995) agree that more sophisticated methods have had limited impact in practice, but still believe that relatively simple methods, such as direct scoring, are applied by designers. However, the widespread application even of such relatively simple methods has been doubted, too (Maffin, 1998).

Are any methods needed?

One may now ask why is it that designers do not apply any methods to support their decision-making processes in conceptual design. Is there no need for applying methods; do designers do very well without them?

Roozenburg and Eekels (1995) claim that many designers have problems with decision-making. They argue, if there are many alternatives and, especially, if many criteria have to be taken into account, designers tend to simplify the problem and apply rules of thumb (heuristic decision rules). This almost always leads to a disregard of relevant information on the decision. Thus, Roozenburg and Eekels conclude, there is a need for methods to help designers make better decisions. This view is shared by other researchers (Ullman & D'Ambrosio, 1995; Wallace & Burgess, 1995).

What are the requirements for support?

So far it could be established that a variety of decision-making methods are available, but apparently none of them is widely used by designers, although there seems to be a need to do so. Are designers not aware of the methods or are the methods not effective?

It is believed that the available methods, neither the mutli-criteria decision-making methods (Ehrlenspiel & Lenk, 1993; Dwaranakanath & Wallace, 1995) nor the methods for modelling uncertainty (Pomerol, 1997) are very effective for conceptual design practitioners. Ullman and D'Ambrosio (1995) agree and claim this is because the methods do not address the requirements for decision-making support in conceptual design. The inevitable question is: w*hat are the requirements for decision-making support in conceptual design?*

In sections 2.2 and 2.3 many researchers were quoted who commented on particular characteristics of the methods under discussion. These comments often assessed characteristics by mapping them onto implicit requirements. However, in many cases these assessments, although being valuable, did not seem to be backed by research, but appeared to be opinions resulting from personal experience. It seems that there are many opinions about method requirements, but little research.

A number of researchers claim that, historically, many methods for supporting design procedures are mainly based on self-monitoring and the subjective experience of talented individual designers rather than on evidence from research projects (Ehrlenspiel & Dylla, 1993; Wallace & Burgess, 1995). Gero and McNeill (1998) claim there is very little literature on how designers design which is based not on anecdotes or on personal introspection. They state that design research over the last three decades has largely concentrated on computer-based models of design. Moreover, Vries et al. (1993) think that the foundations of such models are often not explicit or even wrong, as they were developed without sufficiently understanding the actual design process. The result is that, according to Maffin (1998), current design process models are not widely accepted by practitioners, as these models seldom resemble practical designing. He urges researchers to generate models for interpretation rather than pure prescription. This would provide a logical basis from which to bridge the apparent gap between theory and practice of engineering design and so enable a consensus to emerge between engineering design models and engineering design practice. A need for more research towards increasing the understanding of how design processes actually take place has also been suggested by Shah (1998).

Previous descriptive research

Considering the above I searched the literature for research that aimed at increasing the understanding of how selection-type decision-making in conceptual design actually takes place. Such research, I hoped, would provide insight regarding requirements for decision-making support. In fact, I could identify three studies, which were conducted within the last ten years. These were studies by Ehrlenspiel and Lenk (1993), Ullman and D'Ambrosio (1995), and Dwarakanath (1996) / Dwarakanath and Wallace (1995).

Ehrlenspiel and Lenk (1993) conducted university-based experiments in which students and designers were observed making selection-type decisions. The analysis of these experiments revealed a very rough decision-making process model consisting of five phases (goal-analysis, synthesis, solution-analysis, assessment, result) and a number of interesting observations regarding decision-making processes as they take place. It was not the stated aim of Ehrlenspiel and Lenk to identify particular requirements for decision-making support methods. Their aim was to study the influence of various parameters onto evaluation processes. Thus, although being very valuable, their study has not provided specific answers for the above question on method requirements.

Ullman and D'Ambrosio (1995) as well as Dwarakanath (1996) aimed their studies at identifying requirements for decision-making support in engineering design. Both studies resulted in very interesting sets of decision support requirements (see Appendix B) arranged in frameworks. These results are based on experiments in which *design processes* were observed. This is different from Ehrlenspiel and Lenk who explicitly focused on *selection processes* rather than 'overall' design processes. Because of this difference, I believe, Ehrlenspiel and Lenk's observations regarding selection processes were more specific and detailed than Ullman and D'Ambrosio's and Dwarakanath's observations. Therefore, I suggest, a new study, using Ehrlenspiel and Lenk's focused approach, but having the same aim as Ullman and D'Ambrosio's and Dwarakanath's investigations could potentially further develop the definition of requirements for decision-making methods in conceptual engineering design. In fact, reflecting on his requirements framework, Dwarakanath (1996) himself suggests further studies to improve his framework and thus the foundation upon which to develop effective decision support. Particular findings of these three studies will be addressed in more detail in chapter 5.

Conclusion

As a conclusion from the discussion above I see potential benefit from more empirical research aiming at identifying requirements for decision-making support in conceptual design. It would be worthwhile to adopt a very specific focus on selection-type decision-making processes. Eventually it should be discussed how the characteristics of available decision-making methods relate to requirements as identified through such empirical research. These are the motivations for the study about which the remainder of this thesis reports.

Dwarakanath and Wallace (1995) suggest that, basically, a main reason for undertaking research into decision-making is to develop a greater understanding of the decision-making process. This is seen as a prerequisite for deriving method requirements. Ehrlenspiel and Lenk (1993) suggest an important variable for such research: professional experience. This variable has also been suggested for general research in engineering design by Dixon (1988) as well as Gero and McNeill (1998). They ask: Are there differences between the designing activities of experienced and inexperienced designers? Both suggestions are agreeable to me and are therefore applied.

2.4.4 Research *aim* and *guiding* questions

As evolved in the previous section, a main aim of the research project about which the remainder of this thesis reports is *to identify requirements for selection-type decision-making support in conceptual engineering design.*

As a prerequisite, a greater, detailed understanding of selection-type decision-making processes in conceptual engineering design needs to be developed. A variable to be considered is the 'professional experience of designers': Do decision-making processes of experienced designers differ from those of inexperienced designers?

In chapter 3 on the research methodology I will elaborate on how to gain a 'detailed understanding of selection-type decision-making processes in conceptual engineering design'.

As an orientation for the research, a number of guiding questions can be derived from the various aspects addressed in the discussions on the decision-making methods' specific characteristics, as summarised in section 2.4.1 and 2.4.2. These guiding research questions are listed in the following.

Questions A, B, C and D address the alternative concepts' expected performances with respect to the specific evaluation criteria; questions E, F and G address evaluation criteria and questions H, I, J and K address uncertainty occurring in decision-making situations:

A) Do decision-makers know what range of performance can be expected?

B) Do decision-makers establish effectiveness functions?

C) Do decision-makers assess performances on an independent, absolute basis or relative by comparisons?

D) Do decision-makers assess performances in a qualitative, in a quantitative or in a vague quantitative (fuzzy) manner?

E) Do decision-makers consider differences in criteria-importance?

F) If decision-makers consider differences in criteria-importance, do they assess them on an independent, absolute basis or relative by comparisons?

G) Do decision-makers assess criteria-importance in a crisp or fuzzy manner?

H) Do decision-makers express uncertainty?

I) If decision-makers express uncertainty, what is its nature?

J) If decision-makers express confidence, is it confidence about the likelihood that an alternative will satisfy the criterion or is it confidence about the degree to which satisfaction is achieved?

K) If decision-makers express confidence, how is it expressed?

Eventually, the identified requirements for decision-making support should be compared with the identified methods' characteristics. This is to establish the degree of matching, which would indicate the methods' effectiveness.

2.5 Summary

The literature review has provided a preliminary answer to the question 'how to support selection-type decision-making in conceptual engineering design?' It was mainly a discussion on various available methods. The aim was to identify support methods and to develop an understanding of the field. In particular, an attempt was made to generate insight into the 'manner of thinking and reasoning' i.e. the assumptions and the underlying principles upon which the methods are based.

Through the discussion a number of specific aspects emerged that seem to play a role in methodical decision-making support.

It was then asked whether the methods are actually effective for supporting decision-making in conceptual design. No satisfactory answer to this question could be found. There is an apparent lack of research that aimed at identifying the requirements for decision-making support: without knowing the requirements it is difficult to assess the available methods' effectiveness.

The conclusion from the literature review was that there is a need for more empirical research with the aim to identify requirements for decision-making support. As a prerequisite, a greater, detailed understanding of relevant decision-making processes needs to be developed. To gain an orientation for the research, a number of particular questions were derived from the specific aspects of methodical decision-making support, as found through the main discussion on available methods.

Chapter 3

Research Methodology

3.1 Introduction

After various methods for decision-making support were investigated within the previous chapter, it was argued that there is a need for more research on identifying support requirements. It was concluded that a prerequisite for the identification of support requirements is a detailed understanding of actual decision-making processes. Therefore, the next step in the project upon which this thesis reports was to develop such an understanding. This is the project's main research effort and the results will be its main contribution to the body of knowledge in the field of decision-making support in engineering design.

The aim of this chapter is to present the methodology applied for the project's main research effort. A number of methodological issues will be addressed and for each of these issues alternative viewpoints and techniques will be discussed. These discussions aim at justifying the viewpoints and techniques chosen for my research.

In the following section a general model for stages in the 'greater research process' will be introduced. This provides an overview and, in particular, indicates where my research is positioned within a larger research context. Then, objectives of my research will be addressed. They set the basis for an extensive discussion on the research design. The administration of the data collection is then detailed and followed by a discussion on the limitations of the chosen methodology and ethical considerations. Finally, the last section will summarise this chapter.

3.2 Research process

Research in engineering design should be concerned with understanding design and by using this understanding to change, in particular, to improve the way design is carried out. This requires a theory on the existing situation, a prescription on what is desired and an approach for changing the existing situation into the desired one. (Blessing et al., 1995)

Regarding theories one should distinguish between: (i) data from the existing situation, (ii) generalisations of this data, and (iii) explanations on these generalisations. Neither the data itself nor the generalisations of data are theories – they are simply data or generalisations. Theories are established through explanations on why the data and the generalisations of the data are the way they are. (Dixon, 1988)

Another distinction to keep in mind is the one between descriptive theories and prescriptions. A descriptive theory defines an existing situation and explains why it behaves as it does. A prescription, in contrast, states what is desired.

Descriptive theories may be generated through descriptive studies (Dixon, 1988), which aim at developing an 'understanding' (Blessing et al., 1995) of specific situations. Theories are inductive: they emerge through explanations from a variety of sources and, in particular for complex subjects, they are rarely generated by individuals or small groups, but by many members of the entire research community (Dixon, 1988). This means that individual studies do not necessarily lead to the generation of new descriptive theories. Instead, each study is a *contribution* towards the establishment of such a theory.

Apart from contributing towards theory generation, the results of individual descriptive studies can be used in industry as a basis for guidelines or examples as well as for suggesting what aspects are in need of support (Stauffer et al., 1991; Blessing et al., 1995).

As stated above, if it is the aim to actually improve the design process it is not possible to entirely rely on descriptive studies and descriptive theories – as these concentrate on existing situations only. The development of descriptive theories is essential (Dixon, 1988), but eventually prescriptions are needed based upon the descriptive foundation. This means that the desirable situation has to be modelled, i.e. the improved design process, and it has to be identified how this situation can be achieved. As vehicles for achieving the desired situation, methods may be developed. Prescription is based on assumptions and experience (Blessing et al., 1995). Hence, the transition from having a descriptive theory to suggesting a method based on a prescriptive model is made through assumptions and experience.

The testing of methods is carried out through descriptive studies again (Blessing et al., 1995). These studies describe the design process when the method is being used. By comparing this description with the underlying prescription the method's effect can be assessed and by doing so the method may be validated. Blessing et al. (1995) suggest validation through comparative analysis involving experimental groups (using the method) and control groups (not using the method).

A rather practical research approach implementing the above general considerations to a large extent, has been discussed by Duffy et al. (1995) and is supported by Shah (1998). They suggest a model-method-tool research structure. It commences with a descriptive study and the generation of a descriptive model representing the process as it is carried out. This model is then analysed with the aim to identify perceived process shortcomings. Knowing about the shortcomings of the as-is process, a prescriptive model is directly generated, which is then used for the development of support methods. These methods are implemented as software tools, which practically support designers in carrying out their daily work processes.

The aim of this project's main research effort is to identify method requirements through developing a detailed understanding of actual decision-making processes. From the considerations above it can be derived that such research aim, in particular 'developing an understanding', primarily falls in the realm of descriptive studies and results in a descriptive model. Following Duffy et al.'s (1995) approach, this model may then be analysed to identify process shortcomings. These indicate requirements for support methods as aimed for by my project.

If no data generalisation is involved, any identified requirements for support methods need to be seen as mere *suggestions*, based on only one study rather than a unified descriptive theory. Ideally, the descriptive model would be generalised and used for theory generation before requirements for support methods are identified. Yet, this is not part of my project. As such, the project *contributes* towards theory generation by generating a descriptive model. Apart from this, the project generates specific *suggestions* on requirements for decision-making support methods. The 'greater research process' as outlined above and the main position of this particular research are delineated in figure 3.1.

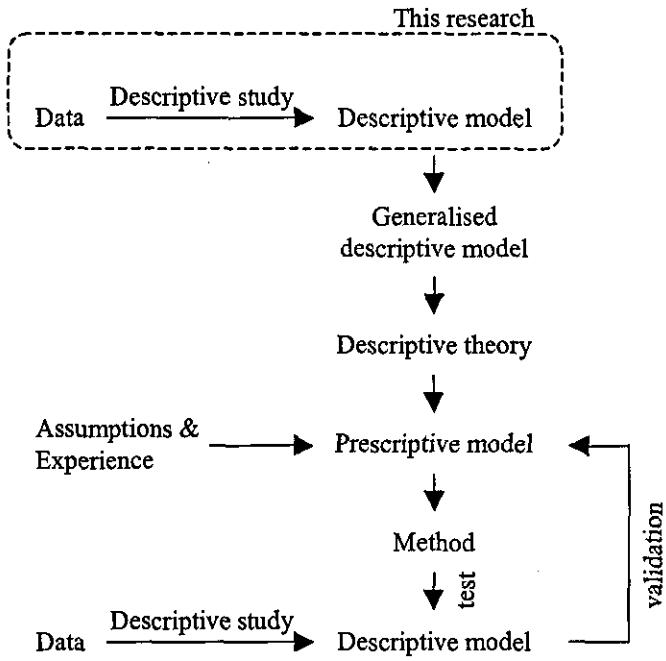


Figure 3.1: This research in a larger research context.

3.3 Specific objectives

Developing an understanding of actual decision-making processes is a research objective, but it is a very general one. Within the previous section it was established that 'understanding' may be developed through generating a descriptive model. This is a more specific research objective: *Generate a descriptive model on how selection-type decision-making processes actually take place in conceptual engineering design.* Yet, there is still the question: how exactly may such a model be generated?

In descriptive design studies, it is typically tried to identify the designers' activities as elements of process models (Stauffer et al. 1991). For example, Stauffer (1987) studied the process of mechanical design; he first identified activities and then used them to model tasks that designers perform.

Models that concentrate on decision-making processes in design usually involve process 'phases' as elements (Tebay, 1984; Ehrlenspiel & Lenk, 1993; Ahn & Dyckhoff, 1997; Feldy, 1997; Pomerol, 1997). These 'phases' are equivalent to the notion of a 'task' (Ahn & Dyckhoff, 1997) as in Stauffer's (1987) study on the design process. Most of the identified previous decision-making process models do not have an apparent descriptive research foundation. An exception is Ehrlenspiel and Lenk's (1993) model. Their model is based on phases and represents a 'rough' process approximation. As *phases* may be modelled by using more elementary *activities* (Stauffer et al., 1991), an approach that first identifies activities and then uses them to model phases will result in a more detailed process model than the one presented by Ehrlenspiel and Lenk (1993). Since my research aims at developing a detailed understanding of decision-making processes I chose an approach for model generation that starts with identifying activities and then uses these activities to model phases. Within this thesis I will refer to process phases as steps because this term seems more simple and therefore more appealing to me. Also, this convention helps distinguishing steps in the decision-making process from phases in the overall design process (e.g. conceptual design phase).

In chapter 2 the variable 'experience' was introduced: It is of interest to identify how decision-making processes of experienced designers differ from those of inexperienced designers. This means that the model must be capable of indicating possible differences. Ehrlenspiel and Lenk (1993) quantified the relative time consumption of the individual decision-making process steps as a metric characterising particular processes. I decided to adopt this metric in a similar form for my research. As I wanted to base my models on elementary activities rather than steps, I had to apply this metric to activities. Hence, I had to measure the identified activities' time consumption as a metric for process comparisons.

From the above considerations I derived the following specific objectives:
Generate a descriptive model of how selection-type decision-making processes actually take place in conceptual engineering design.

- Identify decision-making related activities;
- Quantify the identified activities' relative time consumption;
- Use the identified activities to model process steps.

3.4 Research design

To meet the above objectives I had to design a suitable research approach. Generally, according to the research purpose a particular type of research design is chosen (Robson, 1993). The type of research design may be looked at from different aspects. Once a type of design was clear it needed to be decided which methods should be applied for the collection and analysis of data. There are various methods available. When the methods were selected, a research procedure which uses these methods to meet the research objectives had to be defined. An important issue to be considered for any research design is trustworthiness. Traditional aspects of trustworthiness are validity and reliability.

Within the following sections the design of this research will be discussed. The aims are first, to justify the research design and second, to show openly how I have arrived at my results.

3.4.1 Type of research design: descriptive, exploratory or explanatory

A main purpose of this research is to develop a detailed understanding of decision-making processes. This is needed for identifying requirements of support methods. Developing 'understanding' is a research purpose that falls in the realm of descriptive studies, carrying out descriptive research. This was already established in section 3.2. Traditionally, descriptive research is not about drawing inferences or making generalisations – instead, it is about describing what can be uncovered about some group or population (Mutchnik & Berg, 1996). This type of research has also been defined as gaining "observations with insight" (Leedy, 1989, p.140).

Very similar to descriptive research is exploratory research. The terms 'descriptive research' and 'exploratory research' are sometimes even used interchangeably, although they are not actually entirely the same (Mutchnik & Berg, 1996): the purpose of exploratory research is to understand some group or phenomenon about which little or no previous research has been done; exploration is first required before description may be delivered.

The review of the literature in chapter 2 has indicated that very little previous work has been carried out with the aim to develop a detailed understanding of selection-type decision-making processes in conceptual design. Therefore, an exploratory research design seemed particularly appropriate as the foundation for my research. However, there is still a very descriptive component – the generation of a descriptive process model.

The contrast to exploratory research is explanatory research. The aim of explanatory research is to identify cause-and-effect relationships (Churchill, 1995). Whereas exploratory research is based on research questions, explanatory research is based on testable hypotheses that can be answered with a 'yes' or 'no', or with a precise statement on 'how many' or 'what proportion' (Emory & Cooper, 1991). That is, research questions are open and require words as data to answer, whereas hypotheses are closed and require numbers as data to solve (Perry, 1996).

One of the objectives as stated in section 3.3 asks for the quantification of relative time consumptions. This means that the underlying question actually asks 'what proportion' and clearly requires numbers as data. This could be seen as an indication for explanatory research. Yet, there is no hypothesis involved and the numeric data is not of interest for my research as such. Instead, it is a means for gaining insight about how decision-making processes of experienced designers differ from those of inexperienced designers. 'How-questions' are of an explorative nature (Perry, 1996). Exploring possible differences between groups with different characteristics, such as different levels of experience, has been called 'comparative-exploratory research' (Dwarakanath, 1996).

In section 2.4.4 a number of specific questions were raised 'as a guiding orientation for the research'. This is common practice for exploratory research: the questions should provide an indication for areas of interest, but should not be the only areas explored (Perry, 1996). Therefore I did attempt to find answers to these questions, but they are not the only focus of this research.

3.4.2 Type of research design: qualitative or quantitative

In the previous section the type of research has been looked at from the aspect of overall research purpose: description, exploration or explanation. Another aspect for looking at the type of research is the nature of data required: qualitative or quantitative. The one aspect is not independent from the other – basically they address the same concept: exploratory research is essentially qualitative, whereas explanatory research is essentially quantitative (Perry, 1996). Within this section I will discuss qualitative and quantitative research with the aim to support the discussion in the previous section by looking at the same type of research design from a different aspect.

In general, qualitative researchers prefer narratives and accounts of the way they have interpreted the world, whereas quantitative researchers use mathematical models and statistical tables to relate the research in impersonal terms (Denzin & Lincoln, 1994). Both approaches have particular advantages and disadvantages.

Qualitative approaches are non-numerical. They often produce data that are of considerable depth, but (due to their depth) they are not normally particularly wide-ranged in terms of sample sizes (Mutchnik & Berg, 1996). The qualitative approach is very flexible; it allows for probing and for following different lines of inquiry.

Quantitative approaches are numerical. They allow for producing very wide-ranged data, i.e. data may be collected from a large number of sources (Mutchnik & Berg, 1996). Quantitative approaches require all relevant research questions and all possible answers at the outset of the research (Kennedy, 1997). This makes them inflexible. However, as quantitative approaches do not involve the researcher's qualitative interpretations of meaning, they are little prone to bias.

There has been a debate about which approach to prefer. Qualitative research is criticised because it is seen to lack rigour and objectivity; quantitative research is criticised because it ignores the essential human element of research and is too distant from the subjects of the research (Kennedy, 1997). Yet, Kennedy (1997) argues, such a debate is sterile and instead these two approaches may sit side by side, each informing the other.

Which of the two approaches is most suitable for a particular research should be indicated by the nature of the research question (Mutchnik & Berg, 1996). The research question for this research asks for method requirements and, as a prerequisite, for a detailed understanding of decision-making processes. From the discussion above it can be derived that detailed, in-depth data, can be gained in particular from qualitative approaches. Therefore, this research should be and is mainly qualitative. However, a quantification of time consumptions (see section 3.3) needs numeric data. This particular aspect of the research is quantitative.

Qualitative approaches are generally considered suitable for design research (Kennedy, 1997). A good example is the ground-breaking study by Marples (1960), who investigated decision-tree structures in engineering design. In this study semi-structured interviews and observation techniques were used all generating qualitative rather than quantitative data.

3.4.3 Type of research design: reductionist or constitutive paradigm

A third aspect for looking at the type of research design is its underlying paradigm. A research paradigm is an "overriding viewpoint that shapes ideas and actions about the conduct of research and the validity of its findings" (Cryer, 1999, p.55). As such, reflecting on a research paradigm involves reflecting on the aspects that were discussed in the previous two sections. Choosing a particular paradigm summarises these reflections and takes them a step further by making implicit assumptions and implications explicit.

A point of departure for a discussion on paradigms is the researcher's standpoint regarding subjectivity. This research is to a large extent concerned with generating a descriptive decision-making process model. It was established in section 3.3 that this model is based on decision-making related activities and process steps. These activities and steps are basically identified and defined by me, the researcher, through qualitative, exploratory research. This means quite clearly that there is a substantial amount of subjectivity involved. Obviously, exactly the same applies to any process shortcomings identified through generating the process model and to any derived requirements for support methods.

Subjectivity does not diminish the value of research (Cryer, 1999). However, for an appropriate interpretation of results it is important to acknowledge subjectivity and address it by setting the research within a suitable paradigm. There are two major research paradigms: reductionist and constitutive (Cannon & Leifer, 1999).

Reductionist paradigm

In the reductionist paradigm every phenomenon has a single unique and true description. In other words: all existing descriptions of a phenomenon may be reduced to the one true description (Cannon & Leifer, 1999). This reflects the traditional view on science within engineering and the natural sciences (Eide, 1977). In the literature there are a number of terms for paradigms that are broadly similar to the reductionist paradigm (Cryer, 1999). Examples are: traditional, experimental, scientific, and positivist.

Constitutive Paradigm

In the constitutive paradigm it is possible to create many useful descriptions for a phenomenon. Different descriptions of a phenomenon constitute our understanding of reality (Cannon & Leifer, 1999). This view has emerged as a reaction to the use of reductionist or similar paradigms in the social sciences (Eide, 1997). In the literature there are a number of terms for paradigms that are broadly similar to the constitutive paradigm. Examples are: non-traditional, descriptive, naturalistic, and interpretivist.

Choosing a suitable paradigm

The above two paradigms' major aspects of concern may be summarised as shown in table 3.1 (after Easterby-Smith et al., 1995; Cannon & Leifer, 1999).

Table 3.1 indicates that for my research, the constitutive paradigm is most suitable. This is because (i) it addresses the perspective of this research, which is designer rather than artefact oriented; (ii) it acknowledges subjectivity and (iii) it supports the focus on developing understanding rather than testing hypotheses. This choice of paradigm is supported by Cannon and Leifer (1999) who suggested that design research which is not artefact, but process focused, should indeed be set within the constitutive research paradigm.

Paradigm	Reductionist	Constitutive
Simple characterisation	Phenomenon has a unique & true description	Phenomenon has many useful descriptions
Epistemological agenda	Objectivity (the validity or truth of a model is independent of any observer, but is rooted in the objects in question)	Subjectivity (the validity of a model must take into account – though can be largely independent of – observer and purposes)
Design research perspective	Artefact-oriented (concerned with prescription, in terms of physical properties & mathematical models, for design processes)	Designer-oriented (concerned with description and prescription, in terms of social, psychological, etc. models, for design processes)
Focus of research	Facts, fundamental laws, testing of formulated hypotheses	Meaning, trying to understand
Methods	Measuring, taking large samples, Generalising	Establishing views on phenomena, Small samples studied in depth

Table 3.1: Reductionist and constitutive research paradigms (after Easterby-Smith et al., 1995; Cannon & Leifer, 1999).

Implication

It has been indicated above that research under the constitutive paradigm does not seek to generalise results as it assumes the results are only valid within the specific context in which the study was carried out. However, this does not mean that the results cannot be compared with the results of other studies. They can, only not in a 'universal' arena, as in generalised quantitative results, but in context (Cannon & Leifer, 1999). There is generally no claim that one model is universally superior to another one, or that it is to be considered as 'the truth'. The results can be further used as a means by which to generate ideas and hypotheses for testing and generalisation at a later date (Kaplan, 1986). This is outside the scope of this research project. Yet, it is the mechanism by which this research feeds in the overall research process as discussed in section 3.2.

3.4.4 Methods for collecting data

There are a number of different methods for collecting data. Each of these methods has strengths and weaknesses that need to be known before it can be applied to a particular research study. This research focuses on studying designers to model their decision-making processes. Designers are human beings. Therefore, I only had to consider those methods that may elicit data on human behaviour. Typical methods for this type of data collection are: observation, questions, retrospective protocols and real-time protocols (Stauffer et al., 1991):

Observations

Data collection by observation is very basic. The researcher simply watches a designer at work and makes notes. This method is particularly useful for dynamic 'real-life situations' because there is no set-up for any equipment and obtrusion can be minimised. The weaknesses are lack of detail in the data and the inclusion of the observer's bias.

There are three different types of observations: unstructured observations; structured observations and participant observations. Unstructured observations are made without an agenda. They are used if the researcher does not know what to expect. Structured observations are made with an agenda. They are used if the researcher is looking for particular behaviour. Participant observations involve the researcher as an actual member of a design team being researched. They are used if the researcher has an intimate knowledge of the domain and needs to discern information that cannot be captured by other observation methods, e.g. a team member's thoughts.

Questions

Questioning methods elicit data in response to query. When using questions the researcher is less dependent on the designer than when using observations. This is because rather than passively watching, the researcher may actively ask exactly those questions which are of particular interest. As the questions are not asked during the design process, this method is not obtrusive. Weaknesses are the influence of memory loss (one can only answer as far as one remembers the 'real' incident) and the identification of perceived behaviour (how the respondents perceive or would like to see their behaviour) rather than actual behaviour.

There are three different types of questioning methods: unstructured interview, structured interview and questionnaire. Unstructured interviews are made without an agenda. They are used when the researcher has some knowledge about the domain and is able to form spontaneous questions around specific topics without seeking completeness. Structured interviews are made with an agenda, which is an established set of questions. They are used when the resulting data needs to be complete. Questionnaires are formally printed sets of questions that can be posted. They are used for surveys when large sample sizes are needed.

Retrospective protocols

A retrospective protocol is a detailed record of aspects of a design project. They are prepared by the designer in retrospection. Basically, designers write down exactly what they did during the actual project. As the protocols are not generated during the design project, this method is not obtrusive. Such protocols may deliver voluminous technically very detailed and accurate data. This is because they are the designers' documentation rather than the researcher's. Weaknesses are the influence of memory loss and the identification of perceived behaviour rather than actual behaviour.

Real-time protocols

A real-time protocol is a detailed record of the designers' behaviour and utterances recorded in real-time during the design process. The protocol contains data that is addressed in the designers' short-term memory. This method generates voluminous highly detailed data, which reflects the actual rather than perceived behaviour. The protocols are usually generated through video taping. If individual designers are studied they are asked to 'think aloud'. The result is a 'verbal/visual' protocol. Weaknesses are that this method can be obtrusive, it is not possible to record particular thoughts and it is more difficult to administer than other data collection methods.

Table 3.2 summarises the methods as well as their merits and weaknesses.

Comparison of methods

Due to their specific strengths and weaknesses the methods are not all equally applicable for a particular research. Aspects to consider when choosing a data

Observations Unstructured observations, structured observations, participant observations	
Strengths: ▪ Can be unobtrusive ▪ Easy to administer	Weaknesses: ▪ Lacks detail ▪ Observer's bias
Questions Unstructured interviews, structured interviews, questionnaires	
Strengths: ▪ Unobtrusive ▪ Addresses specific issues ▪ Easy to administer	Weaknesses: ▪ Influence of memory loss ▪ Identifies perceived behaviour
Retrospective protocols	
Strengths: ▪ Unobtrusive ▪ Generates voluminous data ▪ Easy to administer	Weaknesses: ▪ Influence of memory loss ▪ Identifies perceived behaviour
Real-time protocols	
Strengths: ▪ Generates voluminous highly detailed data ▪ Identifies actual behaviour	Weaknesses: ▪ Can be obtrusive ▪ Particular thoughts cannot be recorded ▪ Difficult to administer

Table 3.2: Methods for collecting data (after Stauffer et al., 1991).

collection method for a particular application are: level of obtrusiveness, ease of administration, ability to focus on specific issues, ability to collect detailed data, introduction of bias, influence of memory loss, and ability to identify actual in contrast to perceived behaviour.

Unobtrusive and easy to administer are observations, questions, and retrospective protocols. This makes them particularly suitable for gathering data from 'real' processes (Stauffer et al., 1991).

Especially the use of observations has been recommended for gathering data about how designers go about making decisions (Sen, 1995). Apart from unstructured and structured observations, participant observation has been applied in ethnographic design studies. Ethnography places the researcher into the field or the natural setting of the phenomena under investigation (Berg, 1989). By doing so, this method allows for gathering deep insight through capturing a designer's thoughts. Yet, ethnography suffers from considerable disadvantages (Stacey & Eckert, 1999): the researcher must invest a large amount of time in the study. This is caused by the need to develop an insider's point of view, which in turn requires the researcher to gain the domain practitioners' skills and perceptions. Apart from this, the observations are necessarily incomplete.

Real-time protocols may be obtrusive and they are difficult to administer. This is the reason why they are not usually applied for gathering 'real' design data. It is normally not possible or accepted to follow designers around, in their working environment, with the video and audio equipment to record what they are doing. Therefore, this method is predominantly used in laboratory settings (Stauffer et al., 1991).

The ability to focus on specific issues is characteristic for questioning methods. Structured interviews and questionnaires are very focused, but give little 'informative view' on the phenomenon under investigation (Haque & Pawar, 1998). On the other hand, unstructured interviews may be completely unfocused, offering no guidance at all. To bridge the gap between structured and unstructured interviews, semi-structured interviews have been applied. They allow for a guided study, but still maintain a broad focus (Marginson, 1998).

The ability to collect voluminous detailed data has been attributed in particular to real-time protocols (Sommer & Sommer, 1980). This is because the verbal/visual protocols document the entire process and allow for repeated analysis. Retrospective protocols are also capable of producing voluminous data with a relatively high level of detail. However, a comparative study has shown that the amount of data acquired from a real-time protocol is roughly twice that of a retrospective one (Bradburn & Stauffer, 1991). On the other extreme, observations gain little detailed data (Stauffer et al., 1991).

The introduction of bias is a problem particularly associated with observations. This is because observations give an account of the phenomenon under investigation purely from the researcher's perspective, which may be influenced by previous studies. It is therefore often recommended to commence such study with 'a blank sheet of paper' (Bryman, 1988; Platt, 1988).

The influence of memory loss needs to be considered when questioning methods are used: the results may heavily rely on the respondents recollections (Marginson, 1998). The same applies for retrospective protocols. In contrast, real-time protocols eliminate this problem as they directly capture and record a designer's dynamic behaviour.

The ability to identify actual behaviour in contrast to perceived behaviour is a particular strength of real-time protocols, but it may also be achieved through observations. Retrospective protocols and questioning techniques are vulnerable to distortion and inaccuracy as, in particular, interviewees tend to express the behaviour they would like to exhibit rather than their actual behaviour (Argyris, 1985).

Conclusion

From the above considerations I concluded that real-time protocols would be most suitable for my research. The reason for this choice was principally the method's outstanding ability to gather highly detailed data. Highly detailed data is needed for a detailed understanding of decision-making processes, which in turn may be developed through detailed process models, which is a main objective of this research.

Another reason for choosing real-time protocols was this method's ability to identify actual rather than perceived behaviour. The behaviour of designers as perceived and expressed by themselves may deny some or all of the shortcomings of their decision-making processes. Yet, for this research it is of tremendous importance to find these shortcomings. This is because from the shortcomings, requirements for support methods may be derived, which is the final aim of this research.

Due to the application of real-time protocols it was unsuitable to collect data in designers' normal working environments (see also section 1.4). Instead, the data collection was carried out under laboratory conditions.

Data collection under laboratory conditions has been criticised for not reflecting normal design situations and the results for not simulating the cognitive processes accurately (Shah, 1998). Inevitably, the participants in a laboratory session have less ownership of the decision-problem than in a real-life situation. Yet, Buchanan (1994) argues that this has relatively little impact on the data as the participants more often than not take these sessions very seriously, get involved and finally resemble 'real' conditions closely. Stauffer et al. (1991) found that under laboratory conditions the vast majority of participants feel little distracted by the recording equipment and work 'as normal'. In fact, a number of previous studies in the field of engineering design have also been based on data collected under laboratory conditions (e.g. Ehrlenspiel & Lenk, 1993; Dwarakanath, 1996; Smith & Leon, 1998). Real-time protocol studies have produced various interesting observations in the past and have been applied to generate descriptive models (Shah, 1998).

A specific advantage of collecting data in a laboratory is the possibility of repeating sessions. This is to study different groups under the same conditions (Smith & Leong, 1998), which is impossible if data is collected from the designers' normal working environment. The possibility of repeating sessions without changing the conditions is particularly appealing for this research because it is an aim to find out how the decision-making processes of experienced designers differ from those of inexperienced designers.

3.4.5 Methods for analysing data

As for collecting data, there are various methods for analysing data. Each of these methods has strengths and weaknesses that need to be known before it can be applied to a particular research. Because I chose real-time protocols as the method for data collection, only those data analysis methods had to be considered which can be applied to data generated by real-time protocols. Typically these are: unstructured observations, structured observations and content analysis (Stauffer et al., *1991*):

Unstructured observations

A very versatile and easy analysis method is unstructured observations. It will yield useful information from almost any data. Basically, the researcher studies the data and infers their meaning. Weaknesses are lack of detail and problems with demonstrating confidence in results.

Structured observations

The confidence in results may be increased by structured observations. A structure may be built by defining how the data should be organised into sections. Similar sections can then be analysed together. Examples of such sections are the phases of a design process. Structuring data into sections allows for focusing the analysis and avoids cognitive overload. This allows for a structured repetition of the analysis on the basis of the defined sections. A weakness is that it is still difficult to demonstrate confidence.

Content analysis

A method that is particularly suitable for analysing detailed protocols is content analysis. The essence of content analysis is to define categories of interest and then assign the syntactic data (as kept in the protocols) to them. Practically, this means, the many words in the protocols are transferred into many fewer categories of meaning. These categories can be tested, e.g. with respect to stability by re-coding or with respect to accuracy by comparing the categories with an established norm, if there is one. Through such tests high confidence in the results may be achieved and demonstrated. The main weakness of this method is the effort it requires: content analysis is very time consuming.

Table 3.3 summarises the above data analysis methods as well as their merits and weaknesses.

Unstructured observations	
Strengths: ▪ Easy to use ▪ Versatile	Weaknesses: ▪ Lacks detail ▪ Confidence cannot be demonstrated
Structured observations	
Strengths: ▪ Provides higher confidence ▪ Allows for repeated analysis	Weaknesses: ▪ Difficult to *demonstrate* confidence
Content analysis	
Strengths: ▪ High confidence achievable ▪ Allows for demonstrating confidence	Weaknesses: ▪ Time consuming

Table 3.3: Methods for analysing data (after Stauffer et al., 1991).

Comparison of methods

A comparison of the methods for analysing data may be based on the level of detail in the achievable results, the level of confidence, and the analysis effort.

The least detailed analysis results from unstructured observations. This method will also deliver results without any indication of confidence. This means that if the data is analysed five times one may possibly gain five different results. However, the advantage of such an analysis is that it takes very little effort. More detailed results may be achieved through structured observations, which allow for focused repetition of analyses. This takes a little more effort than unstructured observation. Yet, it is still difficult to demonstrate confidence. High levels of detail may be achieved through content analysis on protocols. When using this method not only structured repetition of analysis is possible, but also the confidence about the categories can be tested and demonstrated. However, many researchers tend to shy away from this method because of the enormous effort it requires (Stauffer et al, 1991).

Conclusion

From the above considerations I concluded that content analysis would be most suitable for my research. The particular reason for this choice was above all, the methods' foundation in categorising the semantic data. Categorising is appealing for this research as it allows for identifying decision-making activities. This means that the activities carried out by the designers can be used as categories upon which the protocols may be structured. The method also delivers detailed results, which is in accordance with my particular research objectives and it allows for achieving and demonstrating confidence in the results, which can be seen as a general feature of research rigour. The method's time demands are balanced-out in this research by keeping the number of protocols to be analysed low. This is in accordance with the constitutive research paradigm (see section 3.4.3).

Categorisation of semantic data from design process protocols has also been used in a similar fashion by Ullman et al. (1988), Gero and Mc Neill (1998) as well as Nidamarthi (1999). The categories in these studies represent 'design episodes', 'design segments' and 'design activities' respectively.

Further analysis

Through the application of content analysis, I was able to identify decision-making related activity categories. By adding-up the time consumptions for each individual appearance of a category throughout an entire protocol, I found the overall time consumption of a particular category within a protocol. This I repeated for each category and each protocol. By doing so and then expressing the categories' overall time consumptions per protocol in percent, I gained the activities' relative time consumptions, as required by one of my research objectives.

Having identified the activities, it was then possible to generate a descriptive decision-making process model, which is also required by one of my research objectives, using a further method for data analysis: pattern coding. The essence of pattern coding is to group categories to constructs, which are units of meaning (Miles & Huberman, 1994). In this research, categories are decision-making activities and constructs are decision-making process steps. Each step consists of a *sequence of activities* and is characterised by a particular activity pattern. In turn, the *sequence of process steps* represents the required descriptive decision-making process model. The relationship between semantic data as in the protocols, categories and constructs is graphically shown in figure 3.2.

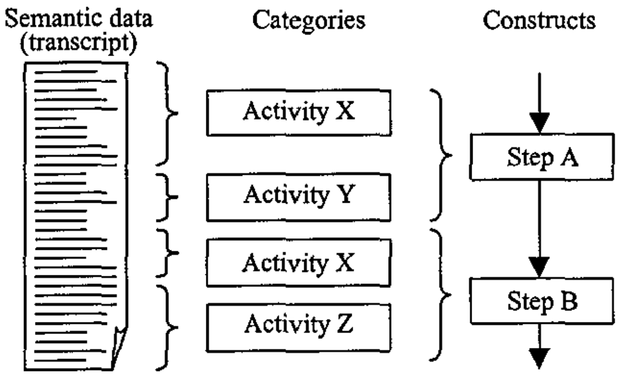


Figure 3.2: Relationship between semantic data, categories and constructs.

Having said above that each step in the decision-making process is characterised by a particular activity pattern it was of interest to analyse whether this really is exactly true. Some steps re-appear a number of times throughout each decision-making process. For these steps I analysed whether the particular activity pattern is exactly the same each time the step appears or whether it is only 'roughly' the same with specific, and possibly very interesting differences. These differences are interesting as they may indicate process shortcomings, as sought by this project. I call such an analysis 'reverse pattern coding'.

3.4.6 Research procedure

So far it has been established that this is exploratory and comparative-exploratory research, mainly qualitative and rooted in the constitutive research paradigm. The data collection is based on real-time protocols generated in a laboratory environment and analysed through content analysis as well as pattern coding.

Research conducted under the constitutive paradigm uses small samples and investigates in-depth. This was shown in table 3.1. For this research, at least two laboratory sessions had to be conducted: one involving experienced designers and one involving inexperienced designers. This was determined by one of the research questions. As a data backup a third session was carried out. Yet, no more sessions than these three were involved in the data collection. This was to ensure an in-depth study within the constraints of this research project's time frame.

Exploratory research, which is of a qualitative nature and rooted in the constitutive research paradigm, may be associated with case study research (Perry, 1996). The main defining characteristic of case study research is the use of a single unit or phenomenon as the focus of investigation (Yin, 1994). The unit may consist of a single individual, a group of individuals, an organisation or a community (Yin, 1994). The phenomenon can be some aspect of human interaction, such as a decision (Marginson, 1998). Hence, this research may be seen as a series of three laboratory based case studies.

Cases

For the data collection I organised workshops taking place in a specially prepared meeting room on campus. The preparation involved setting up recording equipment and taking steps that ensured the workshops were not disturbed. The specific aim of these workshops was to record the participants while they were engaged in making a selection-type decision on a conceptual design alternative.

To resemble a realistic situation I decided to study decision-making groups rather than individuals. This also avoided the need for participants to 'think aloud'. Instead, the naturally occurring conversations could be recorded. For participants I invited final year engineering design students, representing inexperienced designers, and a group of experienced professionals. Students had been 'used' for similar purposes in other studies before (e.g. Ehrlenspiel & Lenk, 1993; Buchanan, 1994).

The participating groups had to complete a given assignment in a way of their choice: I asked them to evaluate a set of conceptual design alternatives and to select, as a group, the one they would suggest for further development. I provided the groups with the alternatives and with some evaluation criteria. All groups had the same alternatives and the same criteria.

As initial pilot studies I ran two workshops which were not analysed. The aim of the pilot studies was to (i) test the recording equipment's effectiveness, (ii) define a suitable group size and (iii) define a time constraint for completing the assignment. This resulted (see also section 3.5) in (i) using a video camera as well as an audio recorder for generating the real-time protocols, (ii) limiting the maximum group size to five members and (iii) set 60 minutes as a time constraint for completing the assignment. This time constraint was a little 'tight', putting a realistic pressure on the participants.

Following the two pilot studies I run three workshops that were analysed. The first one involved a group of three students, the second one involved a group of five students and the third one involved a group of three professionals. In order to supervise the recording equipment I was present throughout each workshop. However, there was no intervention with the groups' decision-making processes except for answering very specific questions. This referred to providing external information, which the groups otherwise would have needed to gain from colleagues, suppliers or the customer. Each of these three workshops was fully recorded.

Transcripts

As a preparation for the data analysis I transcribed all three protocols into spreadsheets (one sheet per protocol) which I had designed for this purpose. These spreadsheets took the form of a table with columns for 'time', 'original conversation', 'abstracted conversation', 'aspect addressed', 'category' and 'format'. The table rows were defined according to the group members' utterances. This means, whenever a participant said something I started a new row, noted the start time in the 'time' column and wrote the exact utterance including a participant identifier, such as 'A: …', in the 'original conversation' column. Some of these utterances had a duration of just one second for example A: *"Yeah…"*. Other utterances could last one minute or even more. I call these utterances 'sections' in accordance with other researchers' terminology (Gero & McNeill, 1998). Occasionally, I also noted the meaning of a participant's gestures, for example *B: "There is nothing…"* (indicating: no problem). The remaining columns were only used for the data analysis. Sections that did not seem to have any relation to the decision-making process were not transcribed. An example is A: *"I'll just get myself a coffee"*. They were timed and noted as 'BREAK'. Irrelevant sections have been treated in a similar way in other studies (Nidamarthi, 1999).

Content analysis

The content analysis, which aimed at identifying activity categories, started with 'a blank sheet of paper'. This means that I did not use any predefined categories, but they were allowed to evolve freely during the analysis.

The first step towards identifying activity categories was to abstract the sections. These abstractions describe the original utterance including which alternative and/or criterion was addressed, if any. For example, I abstracted the utterance C: *"We were assuming the 'weight to be handled' includes the apparatus and the cylinder"* as: Deliberates the meaning of criterion 'weight to be handled'. These abstractions already indicated activity categories. The one above was assigned the category 'definition of criteria'. Thus, the abstractions were intermediate steps towards the identification of categories. In some cases the abstractions revealed that, in particular, longer utterances seem to contain two or more activities in a sequence. Such utterances were then broken down into smaller sections that addressed one activity each. Through rechecking the recorded protocols, each of these sections was then assigned an individual starting time.

The categorisation of sections was not at all straightforward, but was of an iterative nature. When working through a transcript, I constantly defined and redefined categories until once the bottom of a transcript was reached a complete set of categories had evolved. This procedure, which has also been described by Gero and McNeill (1998) as well as Nidamarthi (1999), is shown in figure 3.3. Because of these constant changes in the category set, I had to re-test its applicability by starting from the top of the transcript again, checking whether all sections still agree with the evolved category definitions.

A number of activity categories could also be disintegrated into various sub-categories addressing more specific facets than the category's general description.

Confidence

For gaining increased confidence in the generated set of categories I tested them in terms of stability and accuracy. Stability may be tested through test-retest efforts and

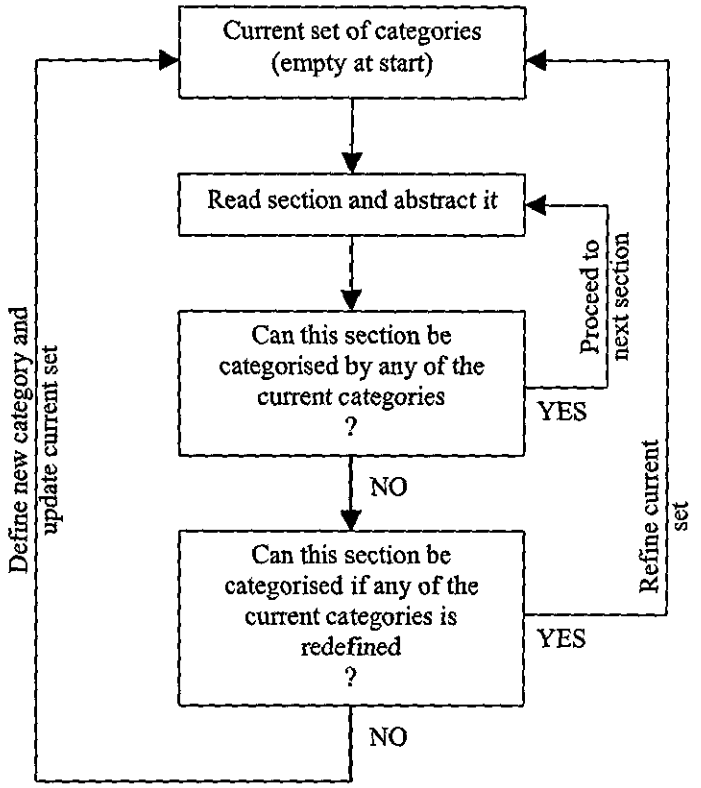


Figure 3.3: Evolution of categories (after Nidamarthi, 1999).

accuracy may be tested through comparing the categories with an established norm (Stauffer et al., 1991).

The evolution of categories through analysing the first transcript was in itself a test-retest approach, as discussed above. Yet, I extended this test across the remaining two transcripts. This means that once a stable set of categories had evolved from analysing the first transcript, I tried to apply this set to the second transcript. By doing so, the set evolved again, which meant that it had to be re-tested for the first transcript as well as for the second transcript. By this means I could arrive at a set which was stable across these two transcripts. This was then extended to the third transcript. At this stage only very few alterations had to be made. Still, once I had worked through the third transcript I re-tested the first and second as well as the third transcript again. By using this so-called replication strategy (Yin, 1994) I could identify a set of activity categories that was generally applicable for all three of my cases. In other words, the set is stable within the confines of this research.

To test the set in terms of accuracy I used a number of previous studies as 'established norm': Ullmann et al. (1988), Dwarakanath (1996), Gero and McNeill (1998) and Nidamarthi (1999). These studies had established a variety of activity categories for the general engineering design process including selection-type decision-making processes, but not focusing on them. To my knowledge, no activity categories have been previously generated specifically with this focus. The effect of differing foci between the previous studies and my research is that my categories are more specific for decision-making processes. However, I could still use the previous studies, at least for a general comparative discussion.

Activities' time consumptions

The use of spreadsheets for transcribing the protocols proved particularly advantageous when it came to determining the different activity categories' time consumptions. The first step was calculating the time increment for each section in the three transcripts. As each section was associated with a particular activity category, the increments could then simply be added-up for each category and each transcript. This could be partly automated through spreadsheet functions, which was very helpful because altogether there were more than 300 pages of transcript. To gain a comparison basis I then expressed the different activities' time consumptions in relative terms, in percent, as to the overall time consumption of each group's decision-making process (the groups did not all take exactly the same time for coming to a decision). Finally, I drew a series of bar charts to visualise the results, which I then compared in a discussion. This gave rise to a number of interesting observations.

Pattern coding

The next phase in the data analysis was pattern coding, which aimed at generating a decision-making process model. This commenced with forming constructs. To do so I studied the abstracted conversations (second column in spreadsheet) again and filtered out process aspects (third column in spreadsheet) that were addressed by the particular sections, i.e. the original conversation (first column in spreadsheet). For example, the performance of concept C2 with respect to the criterion 'time of each delivery' was one of such aspects. Usually, a number of sections, representing a variety of activity categories, addressed the same aspect. I considered such a sequence of activities as a construct. When the aspect changed, for example by the group suddenly addressing the performance of concept C2 with respect to the criterion 'safety', a new construct was formed. By shading different constructs in different colours, the transcripts showed a sequence of constructs as well as a sequence of individual activity categories.

Once I had transformed all three transcripts into sequences of constructs I analysed them with the aim to find terms that described each construct's apparent purpose. For example, the construct addressing the performance of concept C2 with respect to the criterion 'safety' had the apparent purpose of evaluating the concept-criterion pair 'C2-safety'. This purpose became apparent, as the final activity in the construct was to explicitly express an evaluation statement, summarising the activities beforehand. Expressed in abstract terms, the above mentioned purpose became 'evaluate a concept-criterion pair'. I considered this as a process step.

It was also possible to group individual constructs into larger meta-constructs and, in turn, to group these into even larger meta-meta-constructs. The sequence of constructs formed a decision-making process model consisting of various purposeful steps. As I had developed a hierarchy of constructs, the decision-making process model was made of a hierarchy of steps representing meta-meta-constructs, meta-constructs or constructs. An abstract representation of a transcript indicating such construct hierarchy is shown in figure 3.4. Finally, I combined the models developed through analysing each individual transcript into one unified decision-making process model showing different variants of the process. This gave rise to interesting observations.

Inverse pattern coding

Some process steps appeared repeatedly in the generated decision-making process model. These repeated steps were represented in the transcripts by different specific constructs addressing the same abstract purpose. For example, specific constructs with the particular purposes of evaluating the concept-criterion pair 'C2-safety', evaluating the concept-criterion pair 'C2-cost' or evaluating the concept-criterion pair 'C3-ease of use' all address the same abstract purpose, which is to evaluate a concept-criterion pair, and represent therefore the same type of process step. This means, whenever such construct occurred it was the same type of process step being repeated (with a different *specific* purpose).

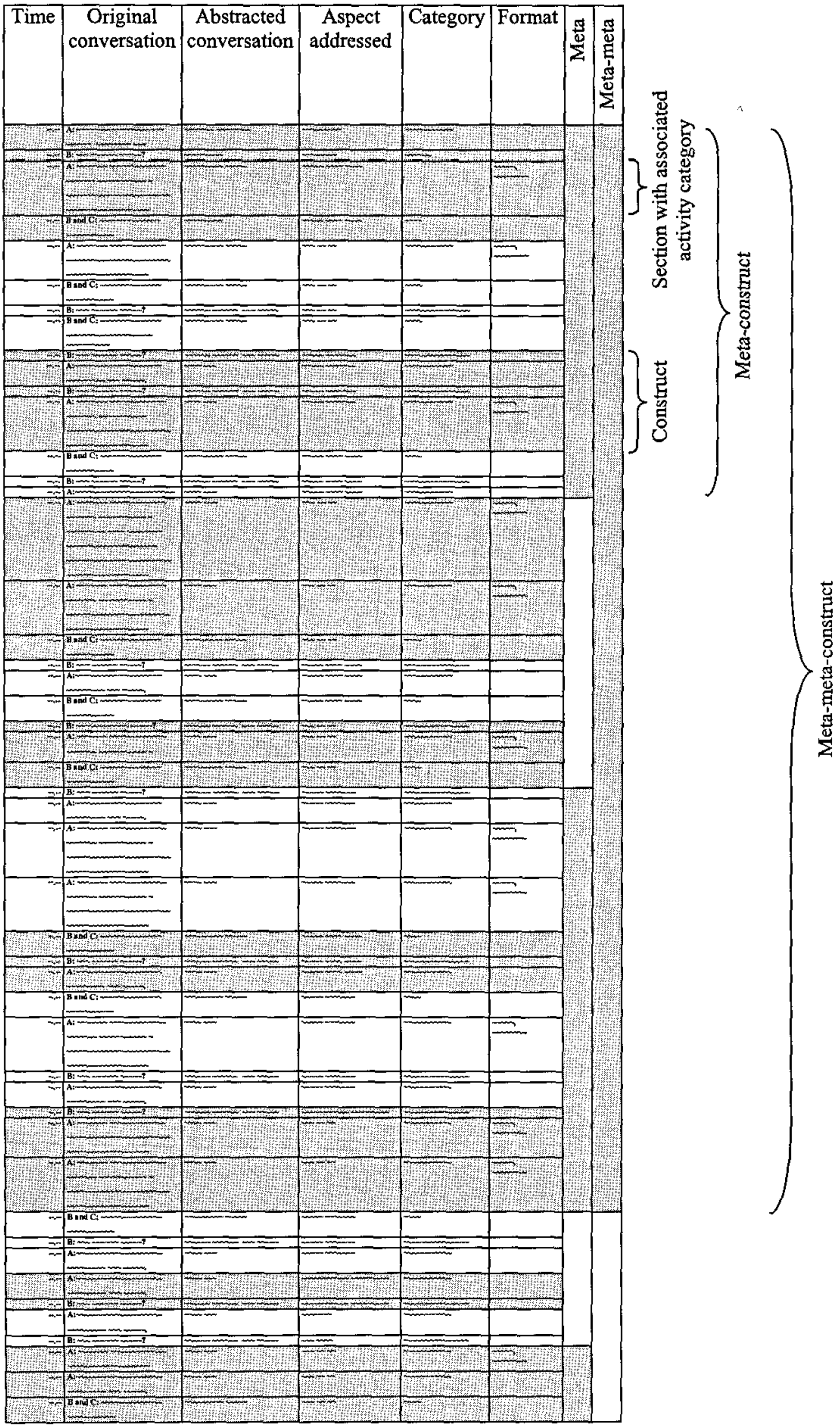


Figure 3.4: Structural representation of transcript with construct hierarchy.

Having identified the repeated process steps I analysed whether the pattern of activity categories in the constructs, which represented the repeated process steps, was always exactly the same or not. Modelling these patterns, which were not always exactly the same, gave rise to interesting observations.

Format of information

To help find answers to the guiding research questions raised in section 2.4.4 I made specific notes on the 'format' of information on criteria importance and of information on the performance of alternatives with respect to criteria. For this reason I had introduced the column 'format' in the spreadsheets. The above classes of information were processed by particular activities (see section 4.3.4 for details). Whenever one of these activities occurred I made a note on the information format. By 'format' I mean, whether it was relative or absolute, qualitative or quantitative, certain or uncertain. If uncertain, was it randomness, belief or vagueness. This gave me clues about possible answers at least for some of the guiding research questions.

There were clues for answering all of the guiding research questions in the transcripts. To find as many as possible I analysed the original conversations again. The identified activity categories and constructs gave directions for this analysis. For example, clues for answering the question on whether decision-makers establish effectiveness functions could be gained by working through the transcripts, focusing on constructs that address the evaluation of alternatives, and see what the decision-makers did. Again, this and the notes on formats gave rise to interesting observations.

3.4.7 Trustworthiness (validity and reliability)

Trustworthiness through validity and reliability is an important aspect of any research design (Mutchnik & Berg, 1996). Validity and reliability are interpreted differently depending on the underlying research paradigm. For qualitative research, set in the constitutive research paradigm, it has been suggested (Perry, 1996) to apply the interpretations of Lincoln and Guba (1985). Thus, their interpretations are the basis for the following discussion that aims at showing how I have catered for trustworthiness within my research design.

Internal validity

Traditionally, internal validity may be defined as a study's ability to plausibly demonstrate a causal relationship between a particular treatment (independent variable) and a particular outcome (dependent variable) (Robson, 1996). The constitutive research paradigm's interpretation, or substitution, of internal validity is *credibility*. Credibility is gained through establishing a match between the constructed realities of the study participants and the reconstructions attributed to them. For my research this is represented by the match between what actually happened in the workshops and the generated models, i.e. descriptions and observations.

The means by which I tried to establish the workshop-model match and thus credibility is called 'persistent observation', which is "to identify those characteristics and elements in the situation that are most relevant to the problem or issues being pursued and focusing on them in detail" (Lincoln & Guba, 1985, p.304). Persistent observation is about providing depth through focus. This has been addressed by my research design in three ways: first, the research focuses specifically on decision-making situations rather than on overall design processes; second, decision-making activities were identified as the relevant elements of descriptive decision-making process models which, in turn, identify the characteristics of such processes and third, the applied data collection and analysis methods have been particularly chosen in the light of gaining detailed understanding, i.e. depth, trading off ease of administration.

External validity

Traditionally, external validity may be defined as a study's ability to produce generalisable results (Robson, 1996). The constitutive research paradigm's interpretation, or substitution, of external validity is *transferability*. Transferability is determined through empirically checking the degree of similarity between sending and receiving contexts. If a study's results are transferred to another context, "the burden of proof for claimed transferability is on the receiver" (Guba & Lincoln, 1989, p.241).

To ease transferability an accurate, 'thick' description of the study is required. Such description aims at enabling someone interested in making a transfer to come to a conclusion about whether this transfer is possible or not. It is not absolutely clear what exactly 'thick' description means, yet it is not the researcher's task to provide any index of transferability, instead it is their "responsibility to provide the data base that makes transferability judgements possible on the part of potential appliers" (Lincoln & Guba, 1985, p.316). As such it was my aim to provide a description as detailed as possible on the research context, in chapters 1 and 2, on the research design, in this chapter, as well as on the research results, in chapter 4.

Reliability

Traditionally, reliability aims at minimising the errors and biases of a study to ensure that the results are independent of the researcher (Marginson, 1998). It is arguable whether this can be achieved at all in case study type qualitative research, set in the constitutive research paradigm, (Marginson, 1998). The constitutive research paradigm's interpretation, or substitution, of reliability is *dependability*. Dependability is concerned with the stability of data (Guba & Lincoln, 1989).

Checking for and ensuring stability was an issue related to increasing my confidence in the activity categories as identified through content analysis (see section 3.4.6). In addition, I also checked these categories for accuracy (see also section 3.4.6) which is another measure to improve confidence and as such relates to reliability as well (Stauffer et al., 1991).

3.5 Administration of data collection

As stated earlier, the data collection took place in a series of workshops on campus of Loughborough University. Before any workshop was actually held I developed a workshop set-up that would ensure an effective and efficient data collection. I tested this set-up in two pilot studies. Taking into account the experiences from the pilots I carried out three main studies to actually collect the required data. The workshops, including the pilots, took place between 15 February 1999 and 5 June 1999.

3.5.1 Workshop set-up

Developing a general workshop set-up involved identifying a specific problem to be resolved by the participants, making room arrangements and preparing a precise agenda.

Specific problem

A design decision-making problem was needed that on the one hand had to be as realistic as possible, but on the other hand had to be simple for practicality reasons. That is, the participants who I had approached beforehand had agreed to spend about one hour, at the most 90 minutes, for a workshop and they were not willing to appear on more than one occasion. Yet, to establish a complete decision-making process model it was necessary to observe the process until it finished with a decision.

A suitable design decision-making problem could be derived from a previous study, which was unrelated to this research: The department of Mechanical Engineering had been approached by a major British industrial company, about an issue related to the delivery of beer-gas cylinders to public houses. The introduction of a new European regulation required changing the current way of delivering these cylinders. The company needed a device which enables one operator to efficiently move beer-gas cylinders from road level into the cellar of public houses and vice versa. The company had provided some criteria for an evaluation of alternative solutions. Within a number of brainstorming sessions involving mostly final-year engineering design students various conceptual solutions for this design issue were generated at Loughborough University. Alternative solutions had also been produced by a group of professional engineers, all having industrial design experience, who did a part-time MSc course at the University. During the development of alternative solutions a number of additional evaluation criteria had also emerged.

The decision-making problem was to select one or a small number of these conceptual solutions for further refinement. For my data collection I reused a sub-set of the solutions. This problem was suitable as it was based on a 'real' design issue and it had the right level of complexity, i.e. it seemed possible to make a decision on one of these solutions within a time period of less than 90 minutes.

For the workshops I prepared handouts which (i) explained the original design issue, (ii) gave a structured list of relevant evaluation criteria and (iii) presented the alternative solutions in the form of sketches and short verbal explanations. I asked the participants to evaluate the alternative solutions, in a way of their choice, within the workshop duration, and to select, as a group, one of them for further development. The complete workshop handout can be found in appendix C.

Room preparations

Preparing a room involved three aspects: (i) identifying a room that was familiar to the workshop participants, (ii) making arrangements that would ensure the workshops were not disturbed and (iii) setting-up required furniture in alignment with recording equipment.

The first two aspects address the attempt to make the conditions of the workshops as realistic and as comfortable (unthreatening) for the participants as possible. The last one addresses the chosen data collection method, i.e. real-time protocols. The group settings required a precise pre-adjustment of the furniture and the recording equipment to make sure that (i) all participants' voices were well captured and (ii) no obtrusive adjustments were needed during the workshops.

Agenda

The following agenda was adhered to during all workshops:

1. The participants made themselves comfortable at a table and became familiar with the aims and methods of this research project as well as with what was expected of them during the workshop. They were also informed about the workshop duration. This was through an informal talk.
2. Through a little presentation using an overhead projector I explained the design issue, showed them the structured list of criteria and introduced them to the alternative design solutions. At the end of this presentation they were given the prepared handouts to be used as reference as well as paper and pens for any notes.
3. The participants started working on their assignment. To observe the recording equipment and to answer questions, without intervening in the group's processes, I stayed in the room throughout the workshop.

4. Once the participants had completed their assignments by announcing the decision I switched off the recording equipment and collected any notes. During an informal discussion I asked the participants to reflect on the process just completed. This gave additional, informal information for the data analysis.
5. The workshops terminated with thanking the participants for their efforts.

3.5.2 Pilot studies

To test the recording equipment's effectiveness, to define a suitable group size and to define a time constraint for step 3 in the agenda, i.e. for completing the assignment, I ran two pilot studies. One of them involved five final year students and the second one involved three final year students.

Initially I planed to exclusively use a video camera for generating the protocols. To visually capture groups between three and five people, the camera needed to be positioned a few meters away from the group. This distance resulted in the audio quality becoming so poor that I found it impossible to generate precise transcripts from such a protocol. The use of a tape recorder with an internal microphone improved the quality, but was still not sufficient. Only when I introduced an external microphone did the audio quality increase to such an extent that it was actually possible to generate precise transcripts. A further, considerable increase in audio quality was achieved when I started using a digital mini-disc recorder in connection with an external omnidirectional microphone. Using this technique instead of conventional audio-tapes resulted in enormous time-savings for transcribing.

It was noticeable that the workshop involving a group of five had a much livelier conversation than the one involving the group of three. The five group members argued about interesting and controversial aspects in great detail. This was positive for my protocols. However, transcribing the five-member workshop was much harder than transcribing the three-member workshop. The reason was that the group of five had much more simultaneous talking than the group of three. Simultaneous talking was very difficult to transcribe and required extensive cross-referencing between audio and video recording. Therefore, I did not involve any groups of more than five members in the main study.

The two groups in the pilot studies took 75 minutes and 40 minutes to complete their assignment. Therefore, I saw 60 minutes as a suitable time constraint for the main workshops. Within this time a decision could be easily made, as shown by the second pilot group, but it would also create a little pressure for slower decision-makers, as the first pilot group.

3.5.3 Main studies

The main workshops which were to be fully analysed involved a group of three final-year engineering design students, a group of five final-year engineering design students and a group of three design-experienced professional engineers, respectively. The participants in each group knew each other well as they had worked in these groups before. I considered this aspect as being realistic which is why I did not reduce the group of five members down to three for easier transcription.

3.6 Limitations of the methodology

The methodology as described in this chapter has one major limitation, which affects the research results. This limitation is lack of generalisability. Basically, generalisability has been traded-off for depth and gaining a detailed understanding of the specific phenomenon under investigation.

The above-mentioned trade-off is rooted in the paradigm underlying this research: the constitutive research paradigm. Through the discussion on research paradigms, in section 3.4.3, but also through the discussions in sections 3.4.1 and 3.4.2 the choice of the constitutive paradigm for this research was justified. In section 3.4.3 it was pointed out that this paradigm acknowledges that research results are only valid within the specific context in which the study was carried out, i.e. results are not generalisabale. Nevertheless, the application of the constitutive paradigm is recommended by mainstream researchers (Cannon & Leifer, 1999) for process oriented design research. Therefore, I consider the lack of generalisability for this research as inevitable and justified.

3.7 Ethical considerations

Observational studies and laboratory sessions do not normally pose risks to the participants (Sommer & Sommer, 1980). However, there is the issue of confidentiality: as the results of exploratory, qualitative studies rely heavily on direct quotations of natural language, participants may become recognisable (Guba & Lincoln, 1989). Within chapter 4 of this thesis I will use many quotations from the transcripts to support my category and construct definitions. To avoid recognition of participants I have not transcribed any identifying information from the protocols. For identifying different speakers I used letters, such as A:... and B:... in the transcripts. For the same reason there are no photographs from the data collection sessions included in the thesis.

3.8 Summary

The aim of this chapter was to present the methodology applied for the project's main research effort. After establishing a set of clear and specific objectives it became evident that this project is a descriptive study leading to the generation of a descriptive decision-making process model. This model consists of decision-making related activities and decision-making process steps. For the empirical identification of these activities and steps an exploratory, qualitative research approach set in the constitutive paradigm was designed. The main research methods involved were real-time protocols for the data collection and content analysis as well as pattern coding for the data analysis.

The data collection was carried out through three workshops in a laboratory setting involving two groups of final year engineering design students and one group of experienced professionals. The groups were given the assignment to evaluate a number of conceptual design solutions and to decide which one should be selected for further refinement. The workshops were recorded and transcribed for the data analysis.

Chapter 4

Results and Analysis

4.1 Introduction

This chapter will introduce the research results, which were gained by analysing the generated real-time protocols through content analysis and pattern coding.

To give a first impression about how the three groups approached their assignment, the following section will provide an outline of their general processes. The aim is to help understanding of the more specific considerations in the subsequent sections. These considerations address the two main results of this research: the identification of decision-making related activities, including the quantification of their time consumptions, and the generation of a detailed decision-making process model. Finally, the chapter will be summarised.

4.2 Process outlines

Three groups of decision-makers were studied. Two of these groups consisted of students and one group consisted of professionals. One of the student groups chose to use a completely informal approach towards their assignment, whereas the other student group chose to use a formal approach. The professionals also chose to use a formal approach.

4.2.1 Students using an informal approach

At the outset, this group took a few minutes (about 5% of their entire decision-making process) to discuss their approach. They agreed on evaluating each alternative with respect to each of the given criteria in turn. Yet, after the first alternative had been evaluated with respect to the first criterion, it was suggested that evaluating the alternative 'as a whole' would be a better strategy. This strategy was then adopted, which meant that each group member expressed what he or she thought about the alternative overall rather than with respect to particular criteria. Subsequently, the group gathered advantages and disadvantages that supported these overall evaluations. Although the advantages and disadvantages could be vaguely related to some of the given criteria (either the more general upper level criteria or the more specific low-level criteria), no criterion was actually mentioned. However, after doing this for some time the group started explicitly referring to the given criteria again and by doing so they reverted back to their initial agreement. In this way, the first alternative's performance was discussed with respect to most, but not all, of the given criteria. In addition, some new criteria were identified. More than 50% of the entire decision-making process' duration was spent in evaluating the first alternative.

About 20% of the process' duration was then spent in evaluating the remaining alternatives. These evaluations were not explicitly based on criteria but, as had been suggested for the first alternative, by first expressing an overall impression and then backing this up through gathering advantages and disadvantages. Again, the gathered advantages and disadvantages could be *related* to criteria.

When all alternatives had been evaluated, the group established a comprehensive ranking. This commenced with informally classifying the alternatives into 'best', 'worst' and 'somewhere in-between'. As they had classified two alternatives as 'the best', a kind of pre-selection took place. The two best alternatives (concepts A and B) were then discussed again and, on the basis of their performance with respect to one particular criterion, comprehensively evaluated in a pair-wise fashion. By this means, a complete ranking was achieved. Eventually, an extensive discussion focused on possible solutions for various identified sub-issues associated with the highest ranked alternative, which was concept A. Through this discussion the group seemed to increase their confidence in this alternative's feasibility. A graphical representation of the group's general decision-making process is given in figure 4.1.

During the entire workshop this group did not take any notes on aspects of their decision-making processes. Within the remainder of the thesis this group will be referred to as the 'informal students'. A more detailed summary of their decision-making process can be found in appendix D.

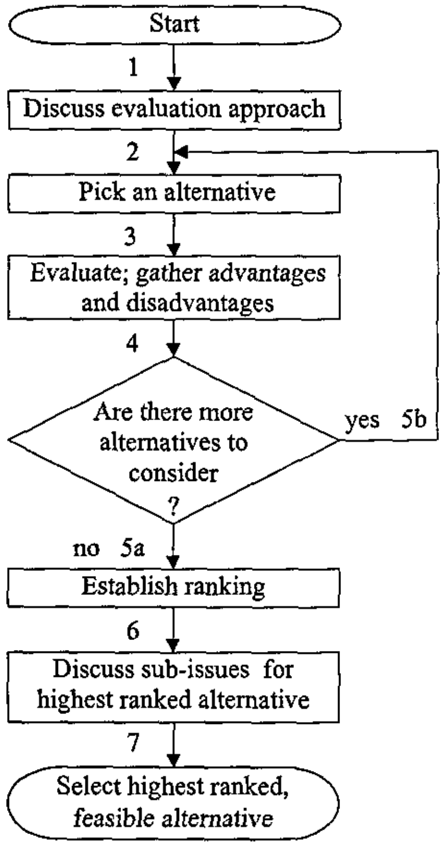


Figure 4.1: The informal students' general decision-making process.

4.2.2 Students using a formal approach

The second student group began tackling their assignment with an elaborate discussion on what approach to use. They deliberated a number of general approaches in detail, but sometimes with little focus. The given criteria structure was altered by cutting off some of the very specific low-level criteria and by distinguishing the remaining criteria into constraints and objectives. This became extensive because for many criteria the meaning and relevance needed clarification and implicit constraints had to be made explicit by reformulating criteria. Altogether, discussing their approach took more than 40% of their entire decision-making process' duration.

Once the approach was clear they drew an evaluation matrix as a basis for documenting their decision-making process and started by assigning weights representing relative importance of the individual objectives (criteria). Subsequently, for each alternative in turn it was checked whether or not the constraints (criteria) were satisfied. Then, for each objective in turn it was scored how well the available alternatives performed in comparison to a datum's performance. The scores were put into the evaluation matrix. As the datum, one of the available alternatives was used (concept A). The scores were then multiplied with the associated importance weights and added-up to generate an overall ranking across all alternatives. The completed matrix is shown in table 4.1. Finally, the ranking was analysed through a group discussion and approved. The group selected concept C2 for further refinement. A

Criteria	Weights	Alternatives											
		A		B1		B2		C1		C2		D	
Constraints													
No lifting of anything over 22kg	c	✓		✓		✓		✓		✓		✓	
No awkward lifting of anything over 15kg	c	✓		✓		✕		✓		✓		✓	
Operator only involved in process	c	✓		✓		✓		✓		✓		✓	
Objectives													
Hazard to staff, customers or general public within the vicinity	5	0	0	0	0	0	0	-1	-5	-1	-5	-2	-10
Efficiency of delivery	4	0	0	0	0	0	0	-1	-4	-1	-4	-2	-8
Initial cost	1	0	0	+1	+1	+1	+1	+2	+2	+2	+2	+1	+1
Running cost	3	0	0	0	0	0	0	+1	+3	+1	+3	0	0
Maintenance cost	3	0	0	+1	+3	+1	+3	+2	+6	+2	+6	0	0
Force required	5	0	0	0	0	-2	-10	0	0	0	0	0	0
Training required	1	0	0	+1	+1	+1	+1	+2	+2	+2	+2	-2	-2
Different cylinders	4	0	0	0	0	0	0	-1	-4	0	0	0	0
Different cellars	5	0	0	-2	-10	-2	-10	0	0	0	0	-1	-5
Σ			0		-5		-15		+5		+9		-14

Key:			
	Weights:	1 not important	5 extremely important	2, 3, 4 intermediates
	Constraints:	✓ satisfied	✕ not satisfied	
	Objectives:	-2 much worse than datum	-1 worse than datum	0 equal to datum
		+2 much better than datum	+1 better than datum	

Table 4.1: Formal students' evaluation matrix.

graphical representation of the group's general decision-making process is shown in figure 4.2.

Through completing the evaluation matrix the group had produced some documentation of their decision-making process. Within the remainder of the thesis this group will be referred to as the 'formal students'. A more detailed summary of their decision-making process can be found in appendix D.

4.2.3 Professionals using a formal approach

The professionals started, as the other groups, by discussing how to approach the assignment. This discussion commenced with identifying and structuring relevant criteria. Their structure was based on the given criteria structure. Yet, a number of alterations were made. They cut off some of the very specific low-level criteria, as the formal students did, and they added some new criteria. The professionals did not distinguish between constraints and objectives. All criteria were treated in the same way and were considered equally important.

Once the criteria were structured, the group started deliberating what methods to use for evaluating the alternatives. They decided to apply a two-stage process. Stage 1 was planned to adopt a relatively crude evaluation method with the aim to pre-select those alternatives that were outstandingly better than the rest. Stage 2 was then defined to adopt a more sophisticated method, leading into a final selection of one alternative. This discussion on the overall approach took about 20% of the entire decision-making process' duration.

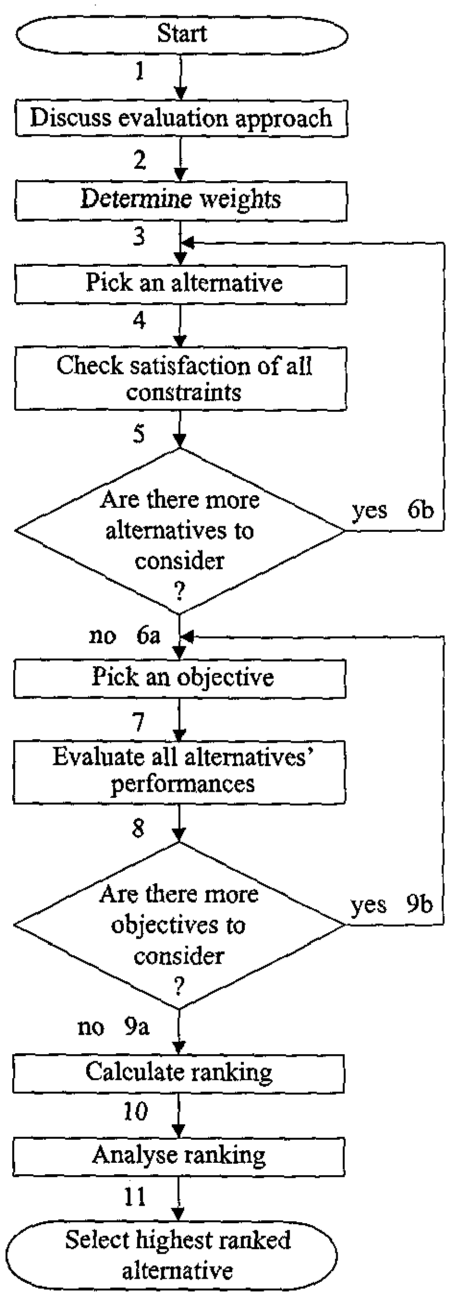


Figure 4.2: The formal students' general decision-making process.

For the first stage the professionals chose to compare the alternatives' performances to a datum. The chosen datum was the manual delivery of the beer-gas cylinders as it is actually carried out. Each alternative in turn was assessed with respect to all evaluation criteria. An evaluation matrix was drawn and the alternatives' performance assessments were put into the matrix as scores. To generate an overall ranking, the scores were added-up for each alternative. The ranking was then analysed which resulted in the pre-selection of two alternatives (concepts B and C2) for further evaluations. The completed matrix is shown in table 4.2.

For the second stage the professionals chose to evaluate the pre-selected alternatives on an absolute basis, i.e. not by comparison. With respect to each criterion in turn, scores were assigned for the two pre-selected alternatives and put into an evaluation matrix. The individual scores were then added-up in order to determine which of the two alternatives scored highest overall. This was concept C2, which was selected for

Criteria	Datum – Current method of delivery	A	B	C1	C2	D
		Alternatives				
Criteria		A	B	C1	C2	D
1 Safety for						
1a) Operator		+	+	+	+	+
1b) Staff		+	+	+	+	+
1c) Customers		=	=	=	=	=
1d) Public		=	=	=	=	=
1e) Cylinder		+	+	+	+	=
1f) Property (new)		+	+	+	+	=
2 Cost						
2a) Initial Cost		2-	-	-	-	-
2b) Running Cost		2-	2-	-	-	-
3 Efficiency of delivery						
3a) Time of each delivery		-	+	-	=	2-
4 Ease of use						
4a) Force required		+	+	+	+	+
4b) Training		-	=	=	=	-
4c) Weight to be handled		2-	-	=	=	-
5 Versatility						
5a) Different cylinders		=	=	=	=	=
5b) Different cellars		=	=	=	=	2-
6 Environmental impact (new)		=	=	=	=	=
Σ-		-8	-4	-3	-2	-8
Σ+		5	6	5	5	+3
Σ		-3	+2	+2	+3	-5
Key:						
+ better than datum	= equal to datum	- worse than datum				

Table 4.2: Professionals' evaluation matrix, stage 1.

further refinement after a brief analysis of stage 2's outcome. The completed matrix for stage 2 is shown in table 4.3. A graphical representation of the group's general decision-making process is shown in figure 4.3.

Through completing their evaluation matrices the group had produced some documentation of their decision-making process. Within the remainder of the thesis this group will be referred to as the 'professionals'. A more detailed summary of their decision-making process can be found in appendix D.

4.2.4 Occurrence of two distinct approaches

As described, two of the groups used distinctively different decision-making approaches to the third group. The difference was made by whether the group used a formal or informal approach. A formal approach was characterised by the application of an evaluation matrix and pre-defined formal scales. This allowed me to consider the chosen decision-making approach as a further variable in my study (in addition to the decision-makers' professional experience).

4.3 Decision-making related activities

By applying content analysis to the three transcripts I could identify a set of activity categories. The confidence in these categories was raised by testing them for stability and accuracy. To identify a metric upon which to compare the three decision-making processes I quantified each category's overall time consumption per transcript. Finally, to find answers towards some of the guiding research questions raised in section 2.4.4 I analysed the possible formats of activity categories addressing the importance of criteria or the performance of alternatives.

4.3.1 Activity categories

The identified set of activities comprises twelve main categories. For a number of activities I could also identify sub-activities, which are differentiated from one another by particular details. In the following explanations each activity will be discussed in turn including any sub-activities. Excerpts from the transcripts are given as examples.

Alternatives	B	C2
Criteria		
1 Safety for		
1a) Operator	1	4
1b) Staff	4	5
1c) Customers	5	5
1d) Public	5	5
1e) Cylinder	3	4
1f) Property	3	4
2 Cost		
2a) Initial Cost	3	2
2b) Running Cost	2	3
3 Efficiency of delivery		
3a) Time of each delivery	3	3
4 Ease of use		
4a) Force required	2	3
4b) Training	3	3
4c) Weight to be handled	2	4
5 Versatility		
5a) Different cylinders	3	3
5b) Different cellars	2	3
6 Environment	4	4
Σ	45	55
Key: Scale: 1 - 5 1 = poor; 5 = excellent		

Table 4.3: Professionals' evaluation matrix, stage 2.

Discussing the process approach

In all cases, the groups discussed how to undertake the actual task of evaluating the alternative design concepts and selecting one. I categorised this activity as 'discussing the process approach'. These discussions addressed either the general approach or the specific approach. Therefore, I divided the category into two sub-categories:

- Discussing the general process approach;
- Discussing the specific process approach.

In discussions on the general process approach the context independent framework of the group's evaluation process was deliberated and determined. For example: *B: "Why don't we go through every design and see whether it meets the criteria first...because then we evaluate every design on the same criteria."*

Within discussions on the specific evaluation approach the context dependent framework of the group's evaluation process was deliberated and determined. The context dependent framework supplements the context independent framework, but is only relevant for the particular concepts and criteria to be involved in the decision-making process at hand. Many of these discussions were about structuring criteria. For example: *A: "We gonna use four headings* [for criteria], *aren't we?" B: "Well, I would use the bottom line ones* [the specific criteria at the lowest level of the given criteria structure] *...not all of them...we could put in a few of our own ones as well, but...I mean the 'safety for operator', 'safety for staff'..." C: "That would go under the general heading of 'safety' then...and then say 1A is...operator..."*

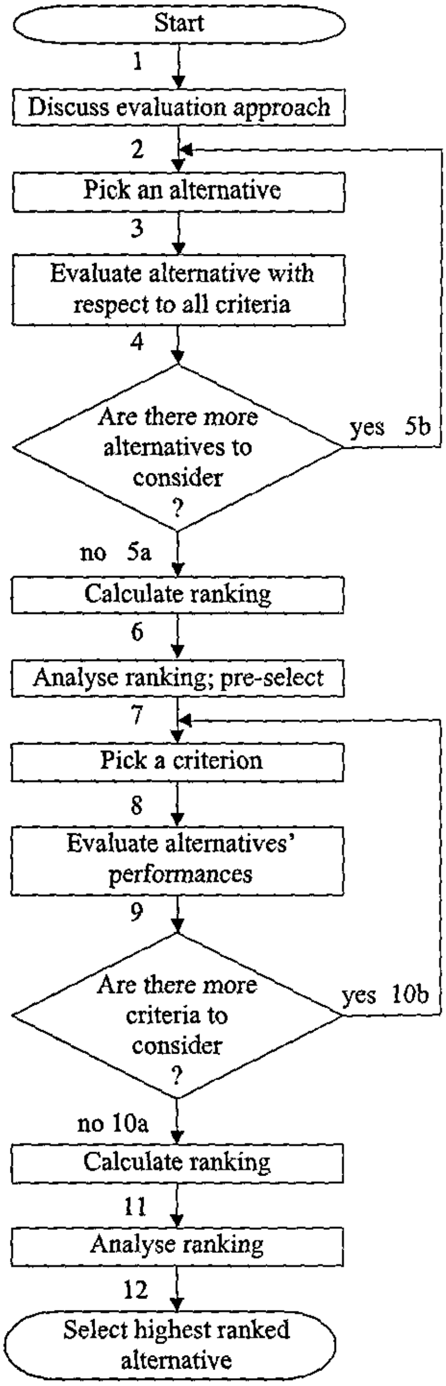


Figure 4.3: The professionals' general decision-making process.

The above two types of discussions predominantly occurred at the beginning of the session in order to organise the overall evaluation activity, but they also appeared at other times during the session either to structure sub-stages, to re-organise the process or to reflect on the process, such as *A: "We are just commenting on the designs...*[looks frustrated]*" B: "Well, we are evaluating the designs, saying which one we go on with as a group".*

Identifying criteria

All groups were provided with a number of structured evaluation criteria. However, during the informal students' and the professionals' decision-making processes additional criteria for the evaluation of alternatives were identified. I categorised this activity as 'identifying criteria'.

The informal students identified criteria implicitly, not intentionally. For example: *C: "It depends on how many* [beer-gas cylinders] *can be pulled up and down."* Most of their newly identified criteria address any of the given higher level criteria. The example above specifies the given higher level criterion 'efficiency of delivery'. The professionals were explicit about the identification of new criteria. For example: *B: "And there should be something about the likelihood of the cylinder itself being damaged...."* This example specifies the given higher level criterion 'safety'. The professionals also introduced a new criterion that did not address any of the given criteria (see table 4.2): *B: "Oh, what environmental things ... is it environmentally ... pollution..." C: "That's a separate section." B: "Yeah, I mean...is it...you could have something that did all of these things* [refers to the other criteria] *perfectly, but it is really noisy and it wouldn't come under...you wouldn't..." C: "Yeah, you're right... 'environmental* [impact]'*..."* (writes it in criteria structure).

Defining criteria

Occasionally, group members indicated that the actual meaning of particular evaluation criteria was not fully understood. The groups then tried to find intelligible definitions of the criterion in question. This seemed to serve the purpose of making sure that the group had a unified and consistent understanding of what each evaluation criterion referred to. I categorised this activity as 'defining criteria'. For example: *B: "... and 'running cost' is including not just the energy used, but also the wear and tear and stuff like that."*

Weighting criteria

In particular the formal students, but also the professionals deliberated whether all the evaluation criteria were actually of the same level of importance. The professionals concluded that they were: *B: "Well, put it in another way. Do we think we need to weight? I mean, maybe they are all important things." A: "Well, they are all important."* In contrast, the formal students did perceive different levels of importance. That is, some criteria were considered more important than others. The perceived levels of importance were communicated by assigning weights. I categorised this activity as 'weighting criteria'. These weights were either expressed informally or formally. Therefore, I divided the category into two sub-categories:

- Weighting criteria informally;
- Weighting criteria formally.

When weighting criteria informally, a criterion's level of importance was expressed without adhering to any predefined format. For example: *C: "I think* [versatility with respect to accepting...] *'different cylinders' is quite important."* Formal criteria weightings did adhere to a predefined format or scale. For example: *C: "Alright, so make it 'five' for* [versatility with respect to being adjustable to...] *'different cellars'."*

Clarifying concept working principles

Occasionally, the working principle of a concept being evaluated, i.e. an alternative solution, did not seem to be fully clear to all group members. If this was realised, the groups tried to gain clarification. I categorised this activity as 'clarifying concept working principles'. For example: *C:* "[addressing concept A] *... the ladder, this telescopic rail, has to actually be delivered with the cylinders now, isn't it? It's not in each pub, this, is it?" B: "Oh yeah, it's something you would ... that's why it's telescopic ... you take it off the back of the lorry and install it."* Clarification was also gained through pointing out major differences between the concepts, such as: *B: "The difference is: this one* [concept A] *has got rails and that one* [concept B] *hasn't. That one needs a barrel role which sometimes is there and sometimes isn't. This one doesn't need anything to be there already. It's a stand-alone design. That's the difference."*

Clarifying the product environment

Unclear environmental conditions were clarified in a similar way to unclear concept working principles. I categorised this activity as 'clarifying the product environment'. Clarifying the product environment aimed at determining the conditions that have a direct influence on the product to be designed and on its performance within this environment. Thus, these conditions needed to be known as a basis for evaluating the alternatives. In cases where discussions revealed that the environmental conditions could not be accurately determined, often due to a lack of information, assumptions were made that enabled the groups to proceed with their decision-making processes. Reflecting this, I divided the category into two sub-categories:

- Determining the product environment;
- Making assumptions on the product environment.

An example for determining the product environment is: *A: "Is this* [the pavement and trap door environment] *usually concrete or is it soil?" C: "This is usually concrete or some brick work ... something solid."* An example for making assumptions on the product environment is: *B: "Well, hang on ... just ... go back ... if it* [a pub's basement] *didn't have a barrel ramp..." C: "But, we are assuming..." A: " ... that it's just a small percentage* [of basements that have no barrel ramp]."

Deliberating sub-issues

The alternative concepts to be evaluated represented solutions to a particular design problem, which I refer to as a design issue. The groups frequently identified and addressed a number of issues that were of a more detailed nature than the stated main issue. These 'sub-issues' were more specific design problems, which would need to be completely resolved at a later stage in the design process. Yet, identifying and resolving these sub-issues at conceptual level often seemed to be necessary for the evaluation of alternative solutions for the main issue. In other words: it was realised that a particular alternative's performance with respect to a specific criterion may heavily depend on means by which specific sub-issues are resolved. Thus, dependency relationships between the main issue and possible sub-issues were identified and deliberated. I categorised this activity as 'deliberating sub-issues'.

Sometimes sub-issues were merely discussed, but at other times solutions for sub-issues were explicitly accepted and assumed to be part of one of the alternatives to be evaluated. Therefore, I divided the category into two sub-categories:

- Discussing sub-issues;
- Accepting assumptions about sub-issue solutions.

In discussions the groups generated and analysed solutions for identified sub-issues. For example, *C: "I was just wondering, there is nothing to stop the cylinders, is there?" E: "Maybe you have to have, say a mat at the bottom." A: "It* [a cylinder] *can be caught by an operator." E: "You think a 22kg bottle falling 6ft can be caught by an operator?" B: "There could be something to hold it..." D: "You could have a little spring at the bottom." B: "Yeah, there is no reason why you shouldn't have a cushion." C: "You would have to."* Occasionally, these discussions were without any apparent outcome. Yet, they could also lead to a result, an assumption, such as: *C: "Yeah, so we are assuming on this one* [concept C2] *that we are going to have some sort of cushion on there.*

Gaining external information

There were situations when the groups required information that was not directly available to them in the workshops. Thus, 'external sources' had to be addressed. These 'external sources' were represented by me, the researcher. This means that the groups explicitly asked me to provide the information sought. I categorised this activity as 'gaining external information'. A situation that required gaining external information was for example: *D:* "[contemplating on the requirement '...removing the need to lift and sustain heavy loads in difficult positions' (see appendix C, product design specification)] *Can you tell us what exactly is a 'heavy load'?" Researcher: "It is 15kg in awkward positions." C: "So, 15kg is maximum." D: "for awkward positions." C: "What is an awkward position?" Researcher: "This is a bit vague, for example bending over is awkward. Lifting a load straight up is not an awkward position. You may then handle up to 22kg."*

Raising evidence

To support different types of statements the groups often tried to raise evidence that could justify either a statement just made or a subsequent statement. I categorised this activity as 'raising evidence'. The supported statements were, in particular, on an alternative concept's performance with respect to a particular criterion (a restricted performance, see following section), on an alternative concept's overall performance (a comprehensive performance, see following section) or on a weighting which assesses the importance of a criterion (see section above). Therefore, I sub-divided the category into three sub-categories:

- Raising evidence on restricted performance of concepts;
- Raising evidence on comprehensive performance of concepts;
- Raising evidence on criteria weights.

Evidence that justifies a statement on a concept's restricted performance was raised for example as: *C: "*[addresses concept C2, criterion 'initial cost'] *...because you have to buy the shoot or the slide ... it's not a great deal, but there is cost there compared to what they are currently doing." B: "But, ... because there are no moving parts ... it's got the potential to be better* [cheaper] *than the previous two* [concepts A and B]*."*

Evidence justifying a statement on a concept's comprehensive performance was raised for example as: *B: "Why has the tube* [concept C1] *come off worse* [comprehensive ranking] *than the slide* [concept C2] *?" E: "Because that* [concept C1] *limits the number we can get down in one go, yeah?" A: "Yeah, and telescope* [concept A] *is low because it's is a bit complicated, isn't it?" B: "It's a lot of hassle..."*

Evidence that justifies an importance weighting for a criterion was raised for example as: *E: "* [deliberates criterion 'weight to be handled'] *If you are doing it all day then you don't want to be pulling lots of weight* [concludes: 'weight to be handled' is therefore a fairly important criterion]*."*

Determining or evaluating performances

When evaluating the different alternatives, the groups communicated expected performances and preferences. I categorised this activity as 'determining or evaluating performances'. There were different types of statements on performances. Generally, it was distinguished between restricted performances (with respect to a particular criterion) and comprehensive performances (overall, with respect to the entire set of criteria). Regarding either type, some statements had an informal character, whereas other statements had a formal character. Sometimes a restricted performance of a concept was informally *determined* without actually *evaluating* it. At other times, the informal wording was chosen in such a way that an evaluation for a restricted performance was expressed explicitly. However, often it was not absolutely clear whether an informal statement regarding an alternative's restricted performance was simply determined, or evaluated. When informal statements were made with respect to the alternatives' *comprehensive* performances, they were always evaluations. This also applied to all formal statements. Therefore, I sub-divided the category into four sub-categories:

- Determining or evaluating restricted performances informally;
- Evaluating restricted performances formally;
- Evaluating comprehensive performances informally;
- Evaluating comprehensive performances formally.

The activity 'determining or evaluating restricted performances informally' was for the informal students the only way of communicating restricted performances and evaluations. In contrast, the formal students and the professionals, who both used evaluation matrices, mapped informal statements onto their predefined formal scales. Such mapping represents the activity 'evaluating restricted performances formally'. For example: *B: " ... alright, B* [concept B] *is better, but not a lot better* [than concept C2 with respect to 'initial cost']. *" A: "Yeah, there is gonna be a slight difference, isn't there?" C: "So, 'initial cost' of B being 3* [formal score] *and the 'initial... ' ... no wait a minute ...it's the other way round, 3 and 2, yeah?"*

The formal students as well as the professionals used their restricted formal evaluations (numeric scores) to calculate a comprehensive evaluation statement for each concept. This is the activity 'evaluating comprehensive performances formally'. Yet, the groups also communicated comprehensive evaluations informally, as did the informal students, such as: *B: "It's* [concept A] *quite good."*

Mapping intuition onto ranking

The formal students as well as the professionals used their formal, comprehensive evaluations to establish ranking orders across all concepts. These calculated rankings were informally analysed by the groups. This analysis basically consisted of statements that represent the mapping of group member's intuitively perceived rankings onto the calculated rankings. That is, the group members expressed whether they felt 'happy' with the calculation result. For example: *B: "So how does that ... take in ... with our feelings?" A: "Contradicts them actually, doesn't it?" B: "Does it?" C: "I don't feel happy with that slide affair* [concept C2]*." B: "I mean I was happy, for instance, that B* [concept B] *was better than A." C: "B better than A? Yes." B: "So, what you don't like is the fact ... so are you happy that B is as good as C1 or do you think it is worse or better?"*

Controlling the process

Group members often made remarks that did not actually contribute towards the content of the decision-making processes, but were aimed at steering the process. As such, these remarks represented process management activities, which I categorised as 'controlling the process'. For example: *B: "Right, let's start with concept A!"*

A summary of all identified activities, as discussed above, is given in table 4.4.

4.3.2 Confidence in categories

The activity categories were identified through cross case content analysis using a replication strategy (Yin, 1994). This resulted in a set of categories being stable across all three of my transcripts. Stability increases confidence in the categories (Stauffer et al., 1991).

Apart from gaining stability I tried to increase confidence through comparing my categories with activity categories identified or observations made by previous studies. As a comparison basis I used the studies of Ullman et al. (1988), Dwarakanath and Wallace (1995) / Dwarakanath (1996), Gero and McNeill (1998) and Nidamarthi et al. (1999) / Nidamarthi (1999). These studies seemed applicable as comparison basis because they had broadly similar research aims. However, these aims were not exactly the same and above all, these studies had applied different research approaches, in particular, different research foci.

ACTIVITIES | Sub-activities	MEANING
DISCUSSING THE PROCESS APPROACH:	The group determines how to work through the assignment of evaluating a number of design alternatives and selecting one.
| Discussing the general process approach	Context independent approach: could be applied to any set of design alternatives and any set of criteria.
| Discussing the specific process approach	Context dependent approach: supplements general process approach and is only relevant for the specific alternatives and criteria at hand.
IDENTIFYING CRITERIA	The group identifies additional evaluation criteria.
DEFINING CRITERIA	The group tries to find a consistent understanding of the evaluation criteria's meaning and relevance.
WEIGHTING CRITERIA	The group expresses the level of importance for evaluation criteria.
| Weighting criteria informally	Using linguistic quantifiers, such as 'quite important'.
| Weighting criteria formally	Using a formal, pre-defined scale.
CLARIFYING CONCEPT WORKING PRINCIPLES	The group makes sure that they understand the working principles of the design alternatives, being evaluated.
CLARIFYING THE PRODUCT ENVIRONMENT	The group makes sure that they understand what the environmental conditions are in which the final product has to operate.
| Determining the product environment	Identifying what the environmental conditions are.
| Making assumptions on the product environment	If the environmental conditions cannot be identified, assumptions are made.
DELIBERATING SUB-ISSUES	The group identifies how an alternative behaves for a particular sub-issue (sub-problem).
| Discussing sub-issues	Generate and analyse solutions for sub-issue.
| Accepting assumptions about sub-issue solutions	Explicitly accept a generated solution for a sub-issue and assume it to be part of one of the design alternatives.
GAINING EXTERNAL INFORMATION	The group gathers information that cannot be generated by them.
RAISING EVIDENCE	The group attempts to justify different types of statements.
| Raising evidence on restricted performance of concepts	Justifying an alternative's performance with respect to a particular evaluation criterion (restricted performance).
| Raising evidence on comprehensive performance of concepts	Justifying an alternative's overall performance. That is, with respect to the entire set of evaluation criteria (comprehensive performance).
| Raising evidence on criteria weights	Justifying a weight for an evaluation criterion.
DETERMINING OR EVALUATING PERFORMANCES	The group finds out how an alternative performs and evaluates the performance.
| Determining or evaluating restricted performances informally	Using informal statements to address an alternative's restricted performance. A perceived value may or may not be communicated.
| Evaluating comprehensive performances informally	Using informal statements to address an alternative's comprehensive performance. A perceived value is communicated.
| Evaluating restricted performances formally	Using a formal, pre-defined, scale to address an alternative's restricted performance. A perceived value is communicated.
| Evaluating comprehensive performances formally	Using a formal, pre-defined, scale to address an alternative's comprehensive performance. A perceived value is communicated.
MAPPING INTUITION ONTO RANKING	The group expresses whether the calculated ranking matches an intuitive ranking.
CONTROLLING THE PROCESS	The group manages (steers) the process without adding content.

Table 4.4: The identified activities and sub-activities.

Ullman et al. (1988) aimed at gaining insight into the way mechanical designs are developed. They generated a task/episode accumulation model (TEA model) through analysing workshop protocols. The workshops had focused on individual designers solving a given problem by producing and evolving design solutions.

Dwarakanath and Wallace (1995) aimed at understanding decision-making processes in engineering design. They analysed protocols of individual designers and groups of designers working on a given design problem. The designers focused on producing design concepts. This study was described in detail by Dwarakanath (1996).

Gero and McNeill (1998) aimed at gaining insight into how designers design. They also analysed protocols of individual designers and groups of designers. The designers had focused on producing different types of designs (electronic and engineering designs) for specific problem statements.

Nidamarthi et al. (1999) aimed at understanding the design process in terms of 'how designers satisfy requirements'. They too analysed protocols of individual designers and groups of designers. The designers were asked to produce concepts for a given design problem, which included a set of requirements to be satisfied. This study was described in detail by Nidamarathi (1999).

As can be seen from the above, all of these studies had similar aims to my study. The closest aim is the one of Dwarakanath and Wallace (1995). Yet, all of these studies focused on design processes rather than specifically on decision-making processes in detail. This is implied by the assignments given to their participants, which was generating design solutions, i.e. *designing*, rather than evaluating and deciding as in my study. This is the reason why it may be expected that the activity categories identified in these studies are not identical to and at the same level of detail as the categories identified in my study. My categories may be seen as specific supplements (regarding decision-making processes) to the more general activity categories of design processes. This means, it should be possible to associate my categories with the more general categories of design processes as identified by the previous studies. Such associations will be discussed in the following with respect to each of my categories in turn.

Discussing the process approach

The activity 'discussing the process approach' was addressed by Dwarakanath and Wallace in their observation 5. They call it 'Argument – Plan'. Gero and McNeill categorised this activity as the micro strategy 'Ds – Referring to Design Strategy'. Ullman et al. called it 'Planning Episode'. Nidamarthi refers, in particular, to the sub-category 'discussing the specific process approach' by his 'Choose' activity, which is about selecting specific requirements or criteria for the design solutions at hand and by his 'Analyse – Relate' activity, which is about establishing relationships between requirements or criteria.

Identifying criteria

The activity 'identifying criteria' was addressed by Nidamarthi in his category 'Identify – Perceive'. Dwarakanath and Wallace have mentioned it in their observation 4. They observed that designers introduce and evolve criteria throughout the design process. This has also been found by Ullman et al.

Defining criteria

The activity 'defining criteria' was addressed by Ullman et al. in one of their observations. They found that designers try to understand their criteria, often fail to do so initially, but resolved this problem later on in the process. Nidamarthi refers to it in his category 'Identify – Modify', which is about altering the definition of requirements or criteria.

Weighting criteria

The activity 'weighting criteria' was addressed by Nidamarthi in his category 'Analyse – Weigh', which is about *informally* expressing the criteria's level of importance. Ullman et al. observed that designers apply a perceived order of importance to their criteria. However, they did not observe that this order was made explicit.

The sub-category 'weighting criteria formally' has not been addressed by any of the previous studies. The apparent reason being that none of these studies involved the application of formal methods, such as evaluation matrices.

Clarifying concept working principles

The activity 'clarifying concept working principles' was addressed by Nidamarthi in his categories 'Evaluate – Relate' and 'Evaluate – Question'. The former is about explaining working principles through analogies and the latter is about identifying unclear aspects. Gero and McNeill have also addressed this activity and categorised it as the micro strategy 'Cl – Clarifying a Solution'.

Clarifying the product environment

The activity 'clarifying the product environment', in particular the sub-category 'determining the product environment' was addressed by Gero and McNeill in their micro strategy 'Ka – Referring to Application Knowledge'. Application knowledge refers to the environment in which the product is to be applied.

The sub-category 'making assumptions on the product environment' has not been explicitly addressed by any of the previous studies. Yet, Dwarakanath and Wallace introduced the category 'Argument – Assumption', which generally acknowledges the occurrence of assumptions in their observed processes.

Deliberating sub-issues

The activity 'deliberating sub-issues' was addressed by Dwarakanath and Wallace in their observation 1. They observed that designers explore sub-issues and generate sub-alternatives before making a decision on the main issue and the main alternatives. Ullman et al. referred to such explorations as 'Sub-Episodes' whose resolution is required by a suspended higher level episode. Gero and McNeill categorised the activity as the micro strategy 'La – Looking Ahead'. Nidamarthi addressed it in his categories 'Generate – Create' and 'Generate – Detail'.

The sub-category 'accepting assumptions about sub-issue solutions' has not been explicitly addressed by any of the previous studies. Yet, as mentioned above, Dwarakanath and Wallace's category 'Argument – Assumption', generally acknowledges the occurrence of assumptions.

Gaining external information

The activity 'gaining external information' was addressed by Dwarakanath and Wallace in their observation 5. They classified external information as information given to designers, such as in handbooks or through researchers' comments. Nidamarthi refers to gaining external information in his category 'Analyse – Question'. Gero and McNeill categorised it as the micro strategy 'Co – Consulting External Information'.

Raising evidence

The activity 'raising evidence' has not been explicitly addressed by any of the previous studies as such. Yet, it may have been covered by for example Gero and McNeill's micro strategy 'An – Analysing a Proposed Solution'. This strategy is about analysing, qualitatively or quantitatively, a solution idea.

Determining or evaluating performances

The activity 'determining or evaluating performances' was addressed by Dwarakanath and Wallace in their observation 5. They call it 'Argument – Product', which refers to *evaluating* performances. Gero and McNeill categorised *determining* performances in their micro strategy 'An – Analysing a Proposed Solution'. *Evaluations* were categorised in their micro strategy 'Ev – Evaluating a Proposed Solution'. Nidamarthi addressed *determining* performances in his category 'Evaluate – Identify Characteristics' and *evaluating restricted* performances in his category 'Evaluate – Verify'. *Evaluating comprehensive* performances with the aim to actually select one alternative was addressed by Nidamarthi in his category 'Select – Compare'. Ullman et al.'s 'Compare Operator' covers both, evaluating restricted as well as comprehensive performances.

The sub-categories 'evaluating restricted performances formally' and 'evaluating comprehensive performances formally' have not been addressed by any of the previous studies. The apparent reason being that none of these studies involved the application of formal methods, such as evaluation matrices.

Mapping intuition onto ranking

The activity 'mapping intuition onto ranking' has not been addressed by any of the previous studies. As this activity is about expressing whether formally calculated rankings match intuitive rankings, it is idiosyncratic for approaches that apply formally calculated rankings. None of the previous studies involved such approaches.

Controlling the process

The activity 'controlling the process' has not been clearly addressed by any of the previous studies. Yet, Dwarakanath and Wallace's categories 'Argument – Process' and 'Argument – Comment' as well as Ullman et al.'s 'Planning Episode' may cover this activity. The closest category from the previous studies is probably Gero and McNeill's micro strategy 'DS – Referring to Design Strategy', which is about designers commenting on the progress of the design process.

Conclusion

The above discussion has shown that most of my activity categories can be associated with categories identified in previous studies. This increases confidence, which is important as these activity categories represent the foundation for the decision-making process models described later in this chapter.

A number of categories could not be associated precisely but still very closely matched the ones of previous studies. These were, in particular, 'discussing the process approach', 'identifying criteria', 'defining criteria', 'clarifying concept working principles' and 'gaining external information'. Other categories, in particular, 'weighting criteria', 'clarifying the product environment', 'deliberating sub-issues' and 'determining or evaluating performances' could be matched to some extent. For 'weighting criteria' and 'determining or evaluating performances' my categories supplement specific categories from the previous studies. These supplements address activities that emerge due to the participants using formal approaches, rather than informal approaches as in the previous studies. For 'clarifying the product environment' and 'deliberating sub-issues' my categories address specific areas of assumptions, which have only been addressed generally in one of the previous studies. Hence, in this respect, my categories specify the categories from the previous studies.

Some of my activity categories have not been clearly addressed by any of the previous studies. These are 'gaining evidence', mapping intuition onto ranking' and 'controlling the process'. Hence, these categories represent new additions to the overall set of categories identified so far.

4.3.3 Time consumption of categories

This section will introduce the activity categories' time consumptions across the three groups' decision-making processes. The time consumptions are expressed in relative terms (in %) to facilitate comparability across the groups and beyond this study. Time consumptions were measured for the main categories as well as for the sub-categories.

Main categories

The main categories' relative time consumptions across the three groups are shown in figure 4.4. The quantity (number of categories) / value (time consumption) distributions in the three charts of figure 4.4 approximate to Pareto distributions. A normal Pareto distribution consists of 'A elements', contributing 20% in quantity and 70% in value; 'B elements', contributing 30% in quantity and 20% in value; and of 'C elements', contributing 50% in quantity and 10% in value. As there are twelve categories, according to a normal Pareto distribution the set of 'A+B elements' (A elements AND 'B elements') would have six categories that together consume 90% of the entire process time. It can be derived from figure 4.4 that this is approximately achieved for each group individually. Five of the six 'A+B elements' are the same categories for each group. These are:

- Discussing the process approach;
- Deliberating sub-issues;
- Raising evidence;
- Determining or evaluating performances;
- Controlling the process.

The charts for the formal students and the professionals also agree with respect to the sixth category in the set of 'A+B elements: 'defining criteria'. This category did, however, not have much time consumption during the informal students' process. This group's sixth category in the set of 'A+B elements' is 'gaining external information'.

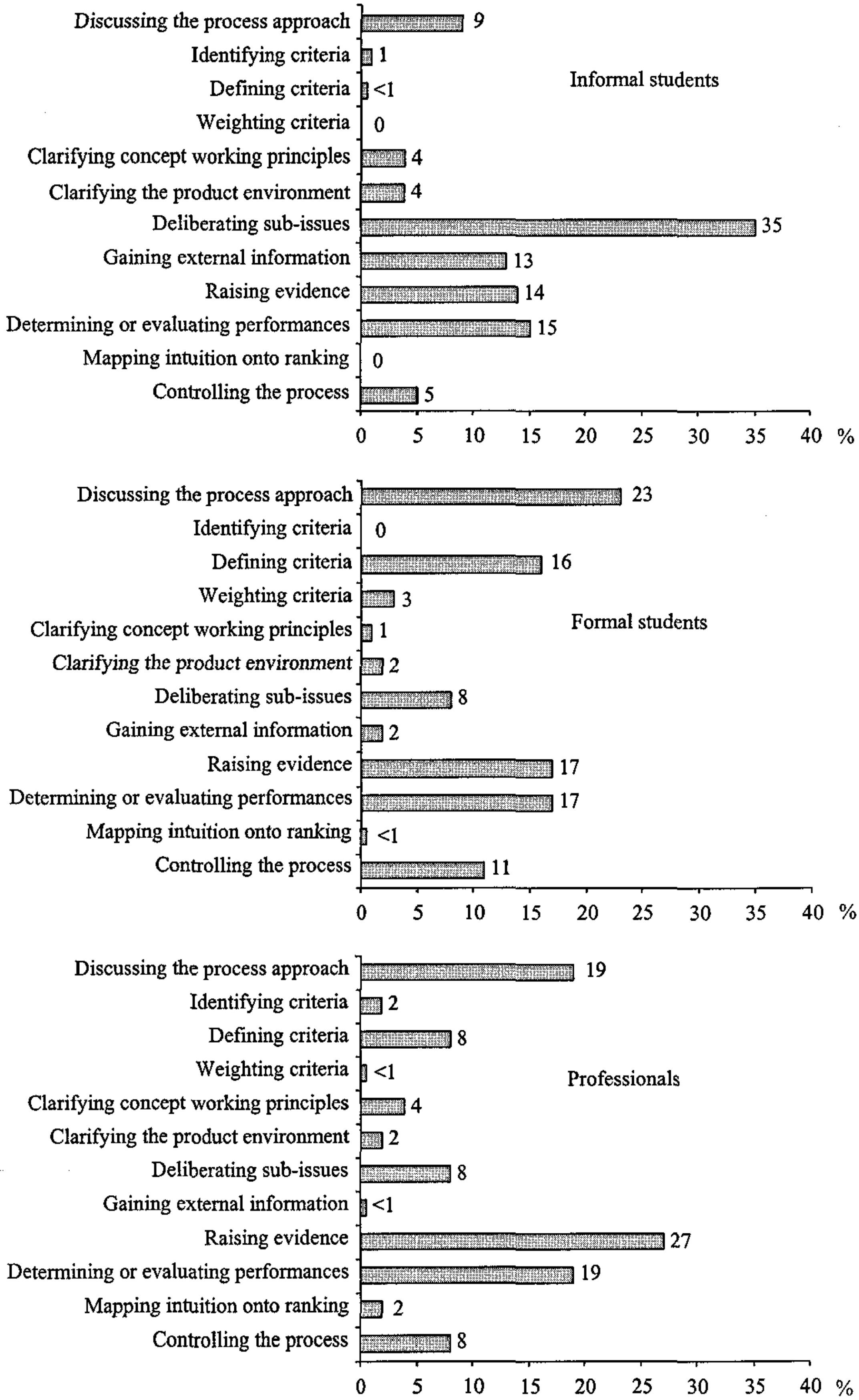


Figure 4.4: Relative time consumptions of the main activity categories.

Even though the charts for the three groups agree on the above five categories as Pareto A+B elements, their individual time consumptions are noticeably different across the three groups' decision-making processes. These differences will be discussed in the following for each of the 'A+B elements' in turn.

Discussing the process approach

The informal students spent a relative time of 9% for 'discussing the process approach'. The formal students and the professionals spent considerably more time for this activity: 23% and 19%, respectively.

Deliberating sub-issues

The informal students spent a relative time of 35% for 'deliberating sub-issues'. The formal students and the professionals spent considerably less time for this activity: 8%.

Raising evidence

The two student groups are rather similar with respect to 'raising evidence'. They spent relative times of 14% and 17% for this activity, respectively. The professionals spent a much higher relative time of 27% for this activity.

Determining or evaluating performances

There is not much difference between the three groups with respect to 'determining or evaluating performances'. The informal students spent a relative time of 15%, the formal students spent a relative time of 17%, and the professionals spent a relative time of 19% for this activity.

Controlling the process

The informal students spent a relative time of 5% for 'controlling the process'. The formal students and the professionals spent higher relative times on this activity: 11% and 8%, respectively.

Defining criteria

The category that only belongs to the set of Pareto A+B elements for the formal students and the professionals is 'defining criteria'. These groups spent a relative time of 16% and 8% for this activity, respectively. In strong contrast, the informal students spent a relative time of less than 1% for this activity.

Gaining external information

The category that only belongs to the set of Pareto A+B elements for the informal students is 'gaining external information'. They spent a relative time of 13% for this activity. Again, in strong contrast, formal students and the professionals spent a relative time of 2% and less than 1% for this activity, respectively.

With respect to the Pareto C elements there are also some distinct differences between the three groups. 'Weighting criteria' and 'mapping intuition onto ranking' are activities that were only carried out by the formal students and the professionals, but not by the informal students. The activity 'identifying criteria' was carried out only by the informal students and the professionals, but not by the formal students. No distinct differences occurred between the groups with respect to 'clarifying concept working principles' and 'clarifying the product environment'.

Sub-categories

The sub-categories' relative time consumptions are shown in figures 4.5 for the informal students, 4.6 for the formal students and 4.7 for the professionals.

Discussing the process approach

There are two sub-categories in the activity 'discussing the process approach': 'discussing the general process approach' and 'discussing the specific process approach'. Whereas the formal students and the professionals spent about equal time shares on each sub-category, the informal students spent twice as much time on 'discussing the general process approach' than on 'discussing the specific process approach'.

Weighting criteria

The activity 'weighting criteria' did not occur at all during the informal students' decision-making process. The professionals did deliberate criteria weights, but it was only addressed very briefly and informally. In contrast, the formal students explicitly weighted the criteria informally and formally. For both of the sub-categories, i.e.

'weighting criteria informally' and 'weighting criteria formally', they spent an equal amount of time.

Clarifying the product environment

There are two sub-categories in the activity 'clarifying the product environment': 'determining the product environment' and 'making assumptions on the product environment'. Whereas the informal students were only concerned with the product environment as far as they could determine it, the formal students as well as the professionals did also make assumptions. They both spent an equal relative time for this activity.

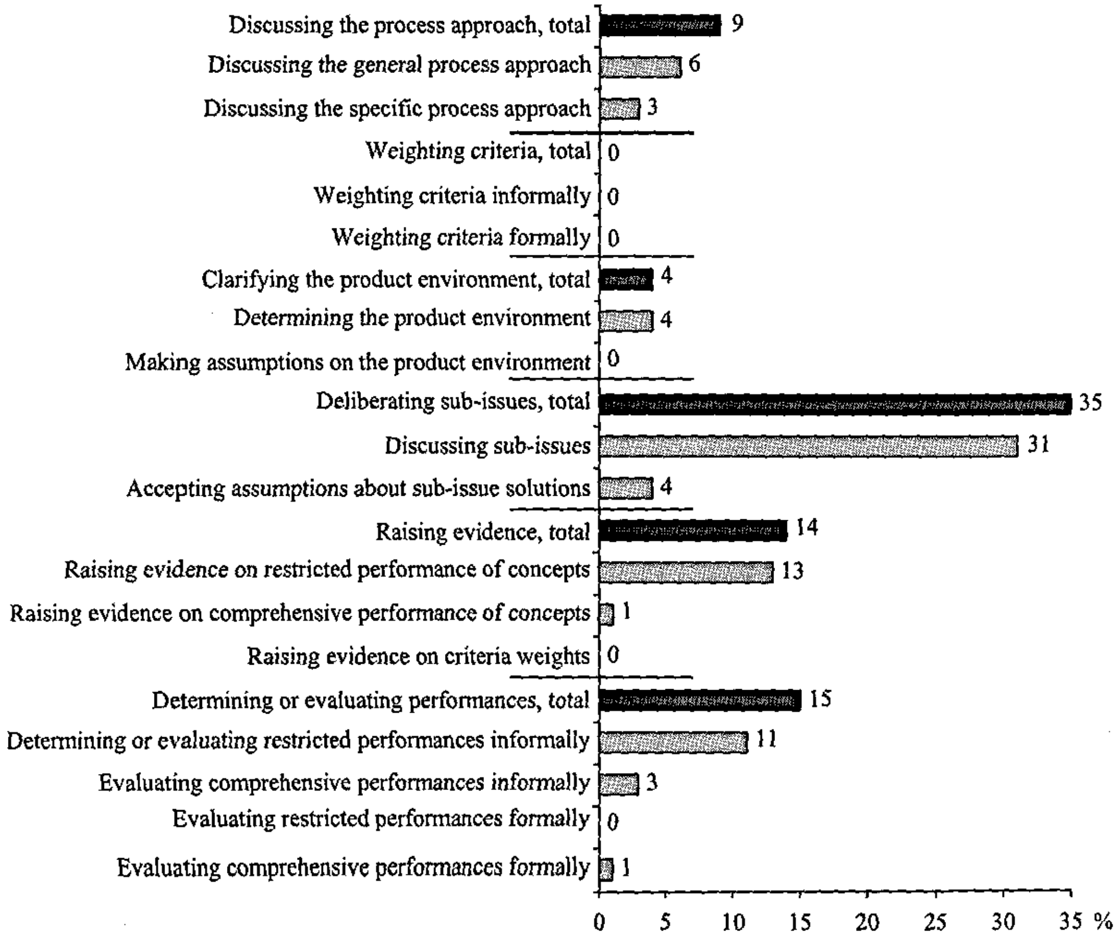


Figure 4.5: Relative time consumption of sub-categories – informal students.

Deliberating sub-issues

There are two sub-categories in the activity 'deliberating sub-issues': 'discussing sub-issues' and 'accepting assumptions on sub-issue solutions'. The former sub-category consumed much more time than the latter. This applies to all three groups.

Raising evidence

There are three sub-categories in the activity 'raising evidence': 'raising evidence on the restricted performance of concepts', 'raising evidence on the comprehensive performance of concepts' and 'raising evidence on criteria weights'. The major time share across these three sub-categories was spent for 'raising evidence on the restricted performance of concepts'. This applies to all three groups. Evidence on criteria weights was only raised by the formal student group.

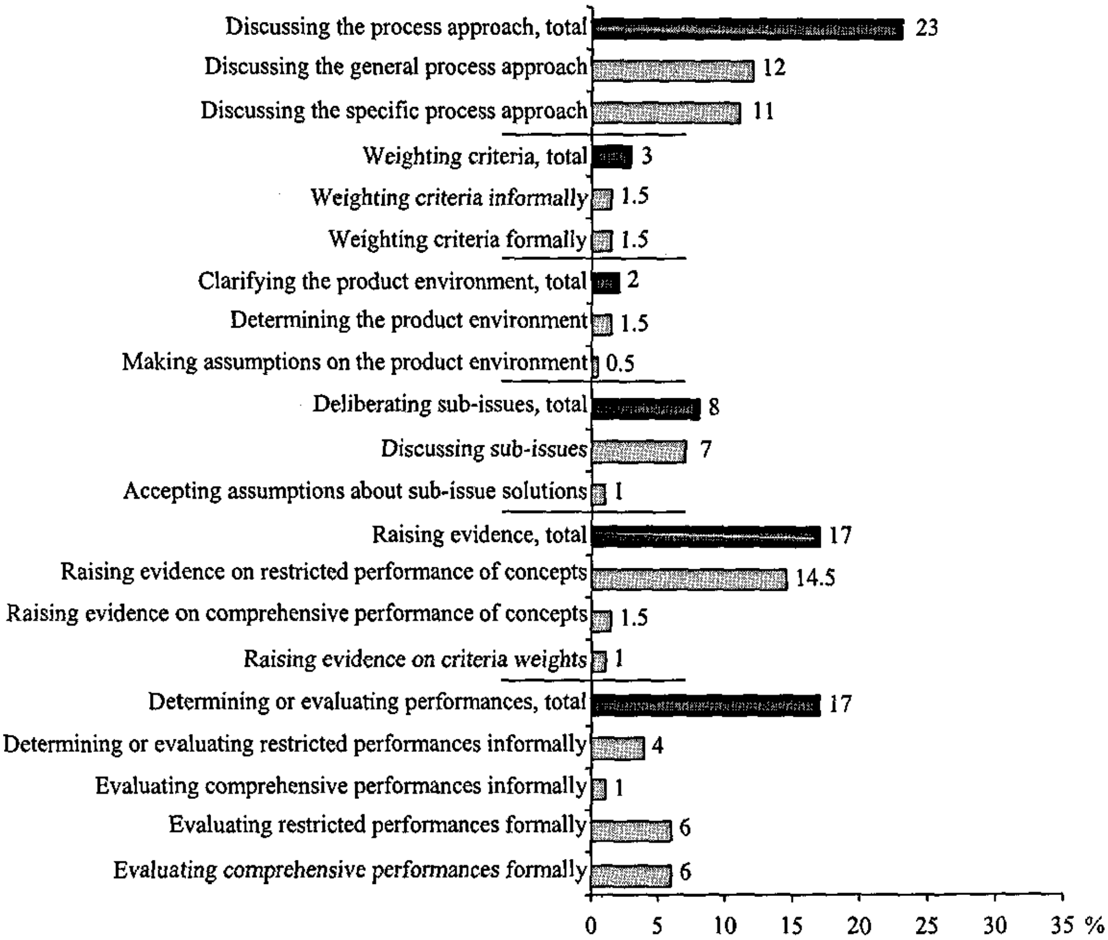


Figure 4.6: Relative time consumption of sub-categories – formal students.

Determining or evaluating performances

There are four sub-categories in the activity 'determining or evaluating performances': 'determining or evaluating restricted performances informally', 'evaluating comprehensive performances informally', 'evaluating restricted performances formally' and 'evaluating comprehensive performances formally'. In the informal students' decision-making process the major time share across these sub-categories was taken by 'determining or evaluating restricted performances informally'. This was different to the formal students' and the professionals' decision-making processes. These groups spent more time on 'evaluating restricted performances formally', an activity that was not carried out at all by the informal students. A distinct difference between the formal students and the professionals is the time consumption of 'evaluating comprehensive performances formally'. The formal students spent a much higher relative time for this sub-category than the professionals.

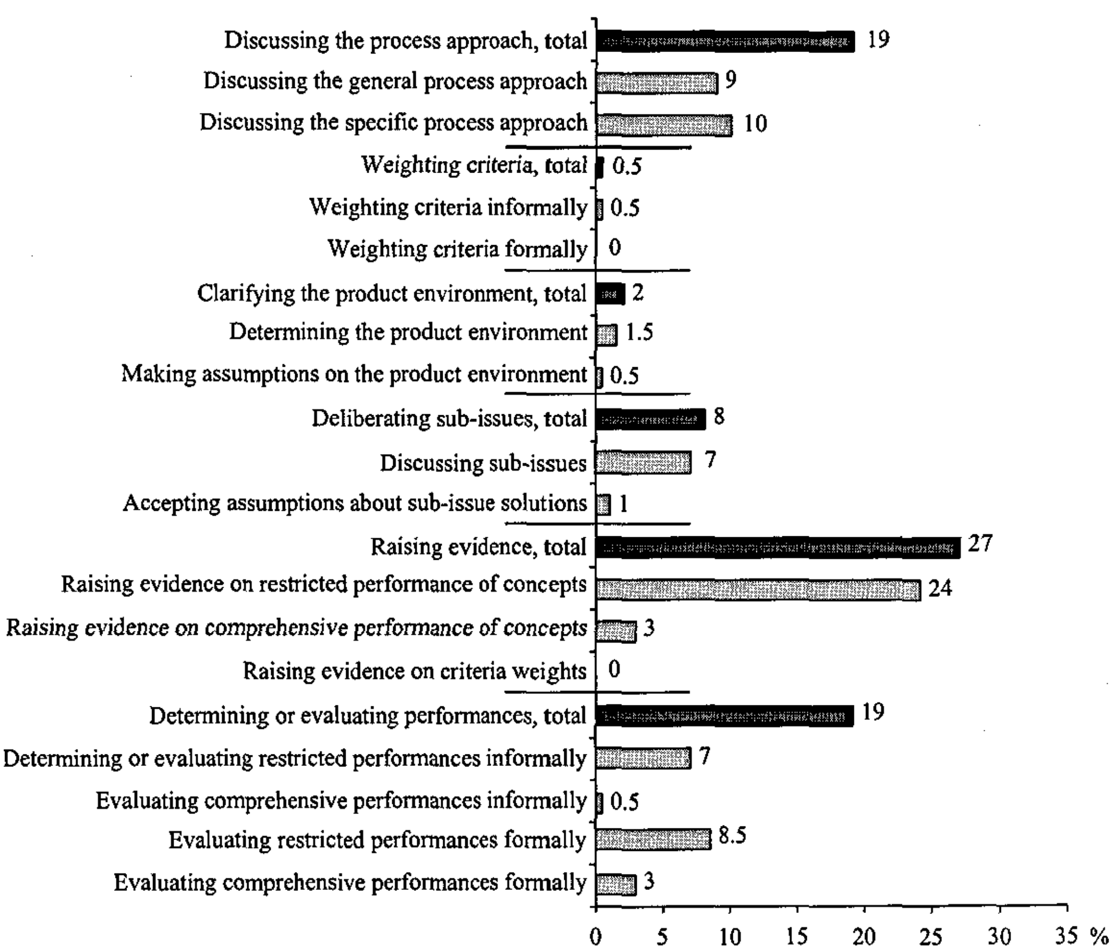


Figure 4.7: Relative time consumption of sub-categories – professionals.

4.3.4 Formats of information

This section will be concerned with the format of information. Specifically, this is information processed by activities expressing the importance of criteria or the performance of alternatives. Relevant activities are 'weighting criteria' and 'determining or evaluating performances'.

Those two groups who used evaluation methods, i.e. the formal students and the professionals, created and then formally forced the formats of information to be put into their evaluation matrices. Information in such 'forced' formats was processed through the activities 'weighting criteria formally' and 'evaluating restricted performances formally'. Yet, I was interested in information formats occurring naturally rather than being forced. Hence, for eliciting such formats I had to concentrate on activities that processed informal information. These were 'weighting criteria informally' for information on the importance of criteria and 'determining or evaluating restricted performances informally' as well as 'evaluating comprehensive performances informally' for information on the performance of alternatives.

Regarding performance information, considerably more time was spent across all three groups for 'determining or evaluating restricted performances informally' than for 'evaluating comprehensive performances informally'. This can be seen in figures 4.5, 4.6 and 4.7. Therefore, I concentrated on the former activity for eliciting formats regarding performance information.

As can be derived from the guiding research questions raised in section 2.4.4, (see questions C, D, F, G, H) three different types of formats were of particular interest: (i) the validity basis, i.e. absolute or relative, (ii) the level of precision, i.e. qualitative, quantitative or vaguely quantified (fuzzy), and (iii) likelihood of information being true (probability/belief). To gain a statement regarding how often information of particular formats appeared in the transcripts, I counted the occurrence of the relevant activities processing information in the formats in question. As the activity 'weighting criteria informally' was only carried out by the formal students to any extent, I used only this group's transcript for counting the occurrence of formats regarding information on criteria-importance.

Validity basis

The validity basis had two formats: either absolute or relative. Absolute may be defined as independent and not related to anything else. For performance information and criteria weights this implies assessments on absolute rating scales. For example, *A: "It would be quick."* (performance) or *D: "That's important!"* (importance of criterion). Relative may be defined as dependent and in relation to something. For performance information and criteria weights this implies assessments through comparison and ranking. For example, *A: "It's safer that way, isn't it?"* (performance) or *E: "I think* [versatility with respect to …] *'different cellars' is more important than* [versatility with respect to …] *'different cylinders'."* (importance of criteria). Figure 4.8 shows the occurrences of the two different formats regarding validity basis across the three groups.

Figure 4.8 indicates a clear distinction between the informal students on the one hand and the formal students as well as the professionals on the other hand. This distinction is with regard to performance information. The informal students predominantly determined or evaluated the alternatives' performances on an absolute basis. In contrast, the formal students as well as the professionals predominantly determined or evaluated the alternatives' performances on a relative basis, i.e. by comparison.

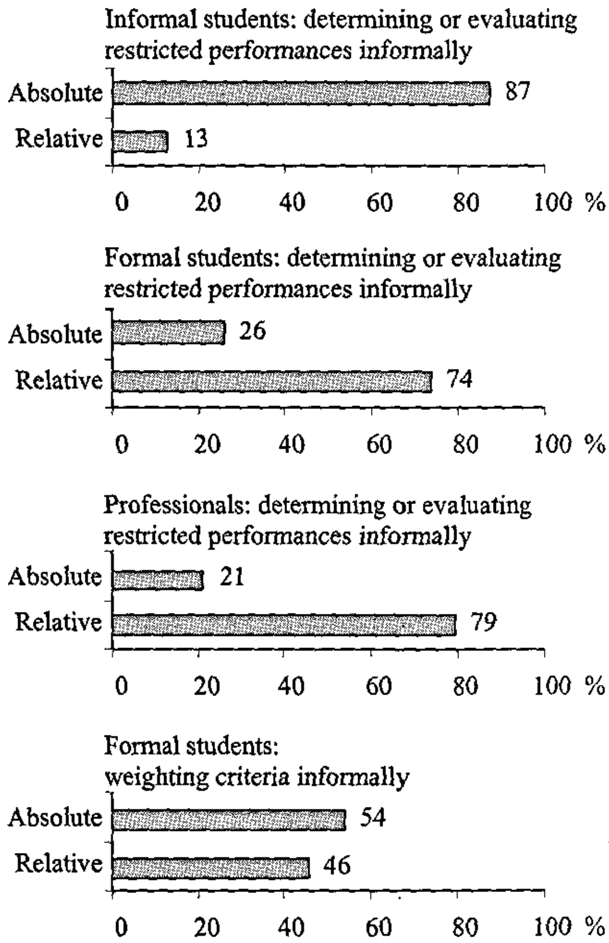


Figure 4.8: Occurrences of the different formats regarding validity basis.

With regard to information on criteria-importance, figure 4.8 indicates there was a balance between weighting criteria on an absolute and on a relative basis.

Level of precision

The level of precision had three formats: qualitative, quantitative or vaguely quantified (fuzzy).

Qualitative information stated whether a particular characteristic was identified or not (binary: yes/no). For example, *A: "Well, it* [concept A] *looks safe..."* (performance, absolute); *C: "There is less prospect of the property being damaged using this* [concept C2] *than the current method."* (performance, relative) or *C: "That's important* [criterion 'efficiency of delivery'] ... *you know, I think this is more important than 'running cost' and 'maintenance cost'."* (importance of criterion, first absolute then relative).

Quantitative information described the extent of a particular characteristic in numeric terms. For example: *C: "It's zero* [concept D and criterion 'running cost'], *isn't it quite honestly ...?"* (performance, absolute) or *A: "... here* [concept A and criterion 'efficiency of delivery'] *he is lowering three* [cylinders] *at a time...." B: "... so it's three times as efficient as*...[an alternative delivering one cylinder at a time]." (performance, relative).

Vaguely quantified information described the extent of a particular characteristic using linguistic quantifiers. For example, *C: "There* [concept A and criterion 'force required'] *is a bit of force required."* (performance, absolute); *A: "It* [concept B and criterion 'force required'] *will be a lot easier* [than current method]." (performance, relative); *C: "I would say 'safety' is quite important."* (importance of criterion, absolute) or *E: "... 'running cost' and 'maintenance cost'...it is a lot more important than 'initial cost'..."*. (importance of criterion, relative). A wide range of these linguistic quantifiers was used by the groups. For instance, regarding the quantification of performance differences: Concept X may be marginally / slightly / a little bit / a fair bit / much / a lot / a great deal better than concept Y.

Figure 4.9 shows the occurrences of the three different formats regarding the information's level of precision across the three groups. The figure indicates quite clearly that with respect to performance information, the prevailing format for the level of precision is qualitative, followed by vague quantifications. Very little performance information had a quantitative format. This applies to all three groups. With respect to information on criteria-importance the prevailing format was vague quantifications, closely followed by the qualitative format. There was no quantitative information.

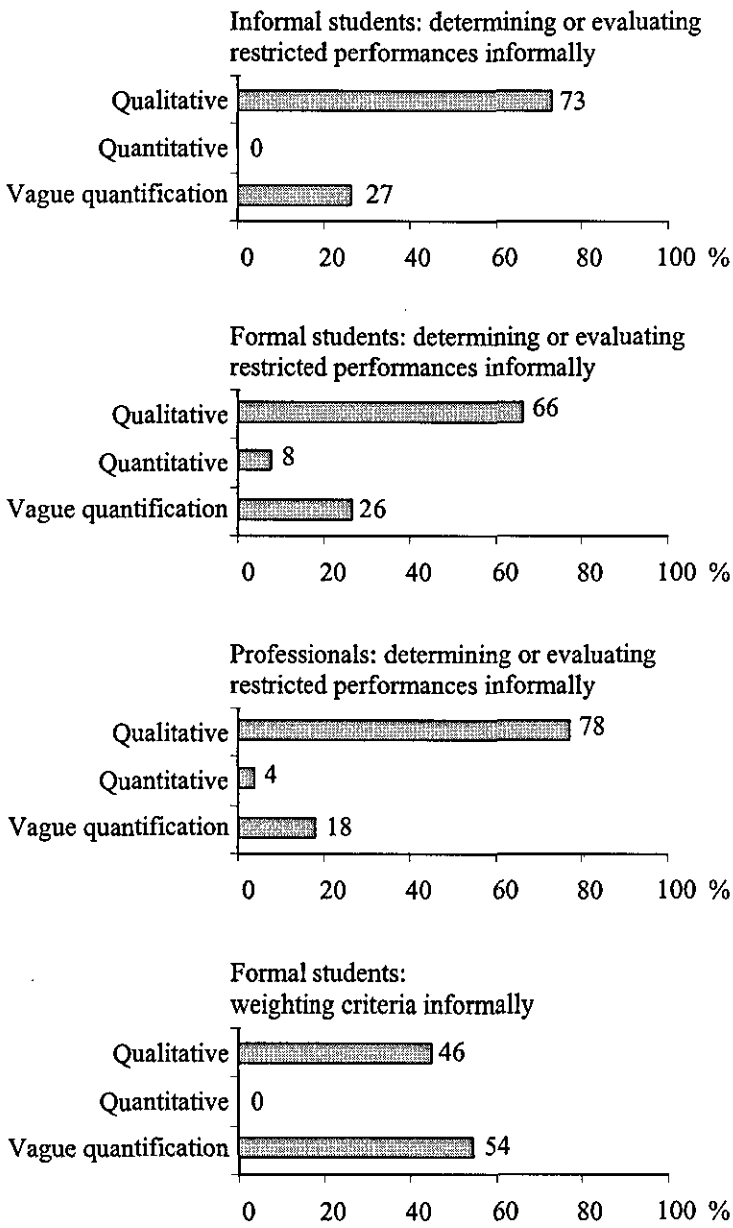


Figure 4.9: Occurrences of the different formats regarding information's level of precision.

Likelihood

Likelihoods were expressed when determining or evaluating performances informally, but not when weighting criteria informally. There were two formats with respect to likelihood. They were either indicated or not indicated. When they were indicated no numeric statements were used, but instead, the groups used linguistic terms. For example: *A: "So yeah,* [if concept A is to be introduced] *there definitely is some training required ... even though it might be just small."* (performance, absolute) or *A: "It* [concept B] *is quicker* [than the current method]*." C: "Yeah." B: " ... as efficient or maybe even quicker." C: "This could be..." A: "... it will be quicker."* (performance, relative).

Figure 4.10 shows the occurrences of the two different formats regarding likelihood across the three groups. It can be seen in this figure that the informal students expressed relatively little likelihoods. Yet, during the formal students' and the professionals' decision-making processes this was different. When determining or evaluating restricted performances informally, these two groups expressed likelihoods much more often than the informal students.

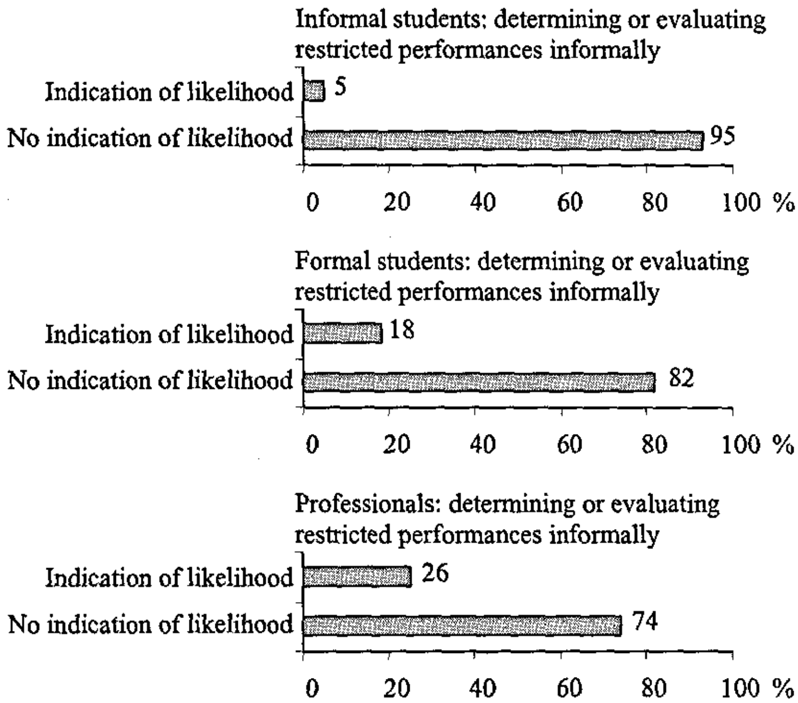


Figure 4.10: Occurrences of the different formats regarding likelihood.

Through further analysing statements containing likelihoods I found that in 65% of all occurrences, across the three groups, a likelihood of 100% is conveyed. For instance: Concept X must be/is definitely better than concept Y. In the remaining 35% of all occurrences a likelihood of 'less than 100%' is conveyed. For instance: Concept X may be / could be / is possibly / is probably better than concept Y.

4.4 Decision-making process model

The activity categories as discussed in the previous section represent the elements of decision-making processes. Yet, since they were considered in isolation it could not be seen how they relate to each other. This section will discuss relationships between activity categories. As a basis for this discussion, a decision-making process model will be developed. The model development commences with identifying the general structure of the three groups' decision-making processes. Subsequently, more specific aspects, which are not relevant for all groups to the same extent, will be considered in detail.

4.4.1 General structure

The general structure of the three groups' decision-making processes comprises three main process steps. These are: step 1, 'prepare the process'; step 2, 'explore criteria and alternatives' and step 3, 'conclude the process'.

Step 1 consisted of a discussion on the general process approach. The groups identified subsequent tasks and prepared to resolve them. Step 2 consisted of investigating relevant evaluation criteria (involving the exploration of their structure and meaning) and evaluating the alternative design concepts (involving the exploration of their performances). Step 3 consisted of ranking the alternatives' overall performances and analysing the ranking. This included aggregating the information gathered during the previous steps and making a final decision to conclude the process. Figure 4.11 shows the general process structure along with the activity categories predominantly occurring during each of the three process steps.

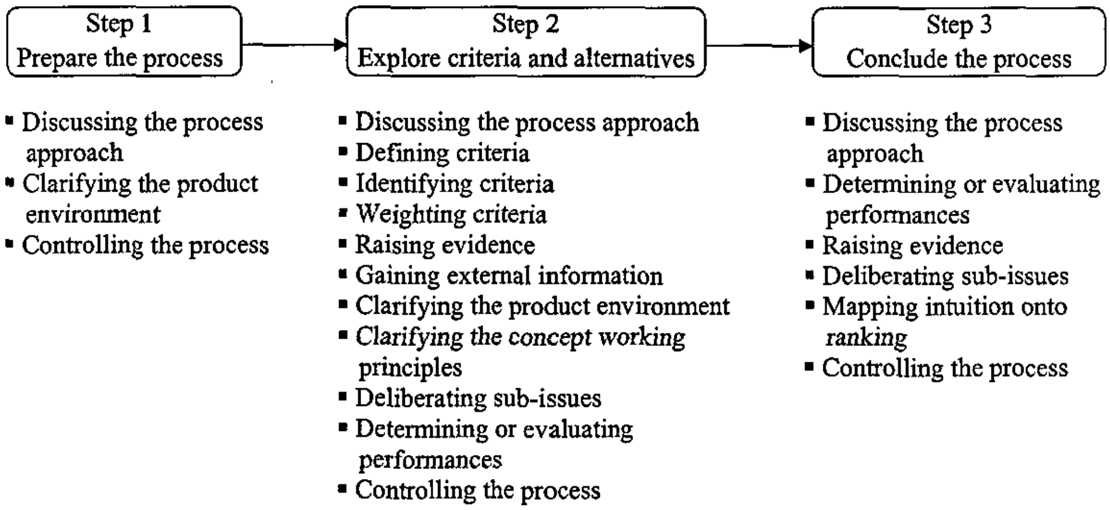


Figure 4.11: General process structure and associated activity categories.

Step 1, prepare the process, involved relatively few activity categories. In contrast, step 2, explore criteria and concepts, as well as step 3, conclude the process, involved a variety of activity categories. Hence, the investigation of relations between categories concentrated on the latter two steps. Appendix D includes summaries of the three groups' decision-making processes and also highlights identified process steps.

4.4.2 Explore criteria and alternatives

There were two different approaches towards step 2, explore criteria and alternatives: (A) comprehensive approach and (B) criteria-based approach. The comprehensive approach was not strictly based on criteria, but considered the alternative concepts 'as a whole', i.e. comprehensively and focused on gathering advantages or disadvantages. In contrast, the criteria-based approach explicitly considered each alternative with respect to each criterion.

Step 2A - Comprehensive approach

The comprehensive approach (2A) was adopted by the informal students. It comprised two steps: 2A.1, 'evaluate alternative comprehensively' and 2A.2, 'support comprehensive evaluation'. In step 2A.1 an alternative was evaluated with respect to its comprehensive (overall) performance. For example, *B: "What do you think of the idea of concept A as a whole?" A: "It's quite good...."* Then, in step 2A.2 the preceding comprehensive evaluation was supported by gathering particular advantages or disadvantages. *For example, A:* " [referring to concept A]... *because one thing ... it's transportable, okay." C: "It's compact, isn't?" A: "Yeah, it's compact." A: "It looks a simple design."* Step 2A.1 seemed in some cases more implicitly 'suggested' rather than explicitly stated. The sequence of these two steps is shown in figure 4.12. It was, however, not strictly and not clearly repeated for all alternatives. It also happened that first the advantages or disadvantages were gathered and then a comprehensive evaluation was made. This is indicated in figure 4.12 by the dashed line.

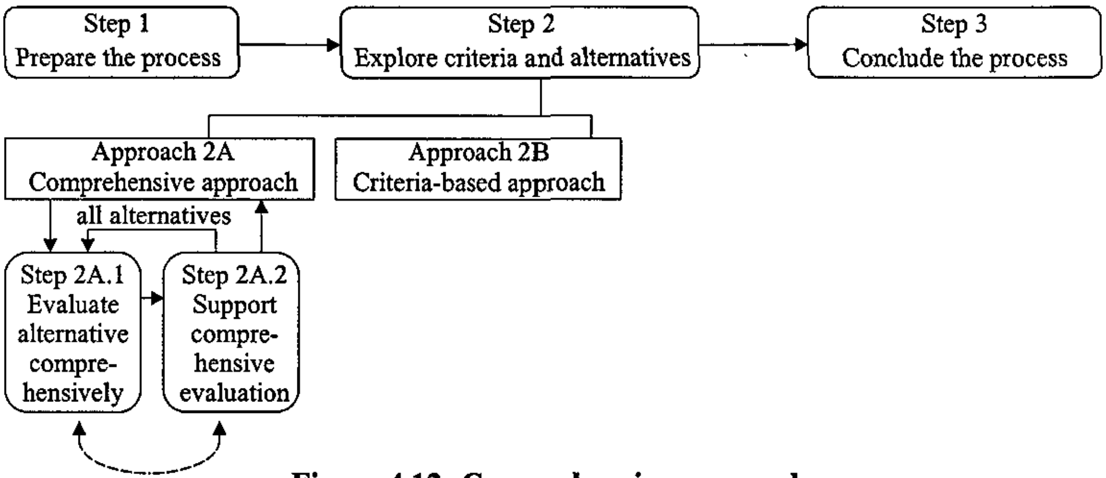


Figure 4.12: Comprehensive approach.

As in the example above, step 2A.1 was an evaluation statement, represented by the single activity 'evaluating comprehensive performances informally'. In contrast, step 2A.2 could involve a variety of activities. These were: 'determining or evaluating restricted performances informally', 'raising evidence on restricted performance of concepts', 'clarifying concept working principles', 'discussing sub-issues', 'accepting assumptions on sub-issue solutions', 'determining the product environment', 'gaining external information' and 'identifying criteria'. The pattern of these activities when carried out as process step 2A.2, i.e. to support one of the informal students' comprehensive evaluations, is shown in figure 4.13.

The activities as shown in figure 4.13 were carried out without any obvious *time* sequence. The activities 'raising evidence on restricted performance of concepts' and 'determining or evaluating restricted performances informally' were always carried out when step 2A.2 occurred. The other activities were not carried out during all occurrences of this process step.

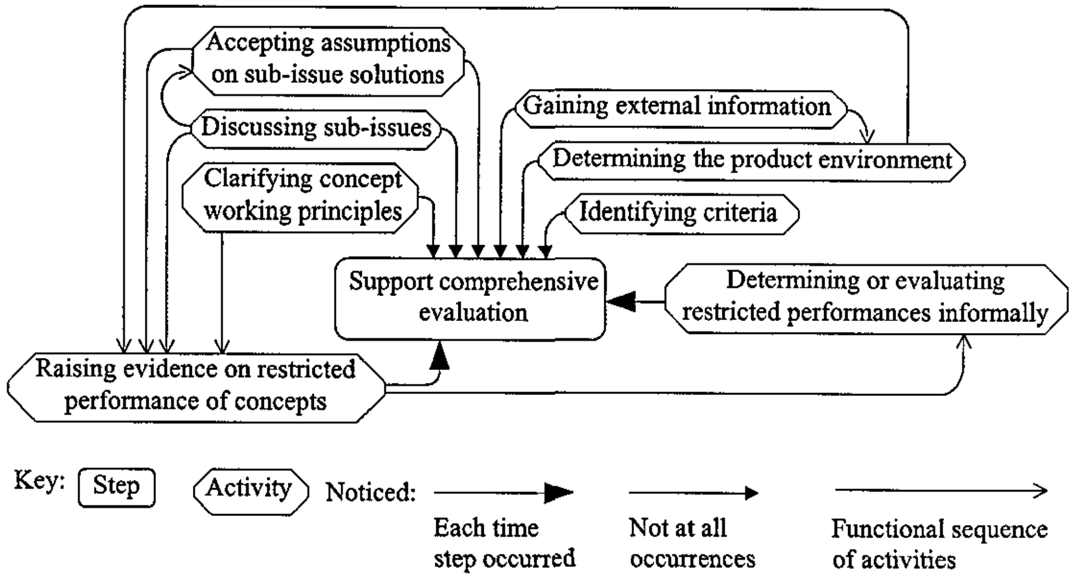


Figure 4.13: Support comprehensive evaluation (step 2A.2).

Through studying the pattern of activities I perceived an apparent *functional* sequence: 'Gaining external information' aided 'determining the product environment'. 'Discussing sub-issues' lead to 'accepting assumptions on sub-issue solutions'. The activities 'clarifying concept working principles', 'discussing sub-issues', 'accepting assumptions on sub-issue solutions', 'making assumptions on the product environment' and 'determining the product environment' all informed the activity 'raising evidence on restricted performance of concepts'. This activity, in turn, delivered a justification for an advantage or disadvantage of the concept at hand. An advantage or disadvantage was expressed through the activity 'determining or evaluating restricted performances informally'. 'Identifying criteria' was not carried out with any perceivable purpose. New criteria simply emerged in the course of resolving this process step. Examples of transcribed activity patterns as modelled by figure 4.13 can be found in appendix E.

Apart from the activity pattern as in figure 4.13 I noticed that the group had a tendency to either concentrate on gathering advantages or on disadvantages. These referred to criteria, which were, however, in many cases not explicitly mentioned (see also appendix D). The set of criteria considered for the evaluation of different alternatives was not constant: different alternatives were evaluated on the basis of different criteria. Table 4.5 shows the criteria that were actually addressed in the different alternatives' evaluations.

Criteria \ Alternatives	A	B	C1	C2	D
Safety	✓		✓	✓	
For operator					
For staff of pub					
For customers of pub					
For general public					
Efficiency of delivery				✓	✓
Set up time (new)	✓	✓			✓
Time for moving cylinders (new)					
Number of cylinders moved simultaneously (new)	✓				
Cost	✓				
Initial cost					
Running cost	✓				
Maintenance cost	✓				
Maintenance intervals					
Requirement of special tools	✓				
Ease of use	✓			✓	
Force required for operation					
Training required for operation					
Weight to be handled					
Ease of handling (new)				✓	
Versatility	✓	✓			
Different cylinders					
Different cellars					
Drop depth	✓				
No barrel role	✓	✓			
Size of trap door					

Table 4.5: Criteria actually addressed by the informal students.

Occasionally, there were iterations or repetitions in the group's evaluations. That is, in particular, concept A was evaluated more than once with respect to specific criteria (see also appendix D).

Step 2B – Criteria-based approach

The criteria-based approach (2B) was consistently adopted by the formal students and by the professionals. It comprised two steps: 2B.1, 'investigate criteria' and 2B.2, 'evaluate alternatives restrictedly'. The evaluation of alternatives in step 2B.2 was based on the criteria structure developed in step 2B.1. It was evaluated how the alternatives performed with respect to (restricted to) each criterion individually. The two steps are shown in figure 4.14.

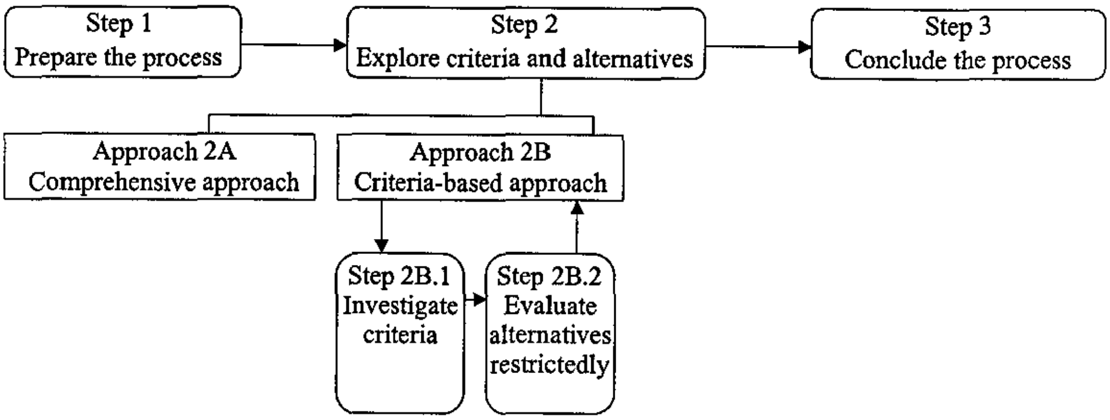


Figure 4.14: Criteria-based approach.

Step 2B.1 - Investigate criteria

The investigation of evaluation criteria in step 2B.1 could be further broken down into three steps: 2B.1.1, 'structure criteria', 2B.1.2, 'pick a criterion' and 2B.1.3, 'determine importance of criterion'. Within the first step the decomposition of the given criteria into sub-criteria was discussed and evolved. Also, additional criteria were identified and the groups tried to clarify the meaning of criteria. Within the subsequent two steps the relative importance of criteria was determined for each criterion in turn. The three steps are shown in figure 4.15.

Step 2B.1.1 - Structure criteria

Structuring criteria in step 2B.1.1 consisted to a large extent of the activities 'discussing the specific process approach' and 'defining criteria'. During this step, the given set of evaluation criteria was re-structured and the individual criteria's definitions were explored and clarified. For these clarifications, external information was occasionally gained. When new criteria were identified, they were defined and incorporated in the criteria structure, which was drawn up as part of an evaluation matrix (see tables 4.1 – 4.3).

The formal students as well as the professionals pruned the criteria structure given to them. The professionals also extended the structure by adding new criteria. Most of the new criteria, such as 'safety – for cylinder', were grouped under given criteria, such as 'safety'. However, the professionals also added a completely independent criterion, which was 'environmental impact'. The addition of 'time for each delivery' was seen by the professionals not as the identification of a new criterion but, as stated by one of the group members, the identification of a metric for the existing criterion 'efficiency of delivery': C: *"So we are saying 'time of each delivery' is the measure of 'efficiency* [of delivery]*'."* The criterion 'efficiency of delivery' had originally not been further decomposed at all.

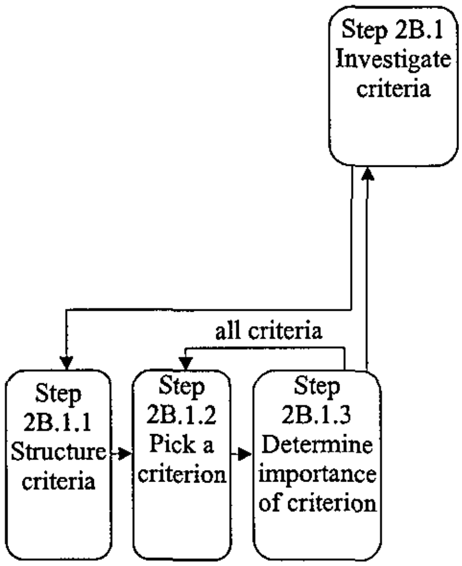


Figure 4.15: Investigate criteria (step 2B.1).

A distinction of criteria into constraints and objectives was made explicit by the formal students, but not by the professionals. Nevertheless, the existence of constraints was implied by the professionals too. This was through utterances like: *C: "...constraints to be met...if we extend the amount of time they* [delivery person] *use on every job, they are not going to bother using equipment...I think, we have got to think about something that says: If it's gonna add a 100% ... if it's gonna double the amount of time, ...if instead of doing the job in 10 or 15 minutes it's gonna be half an hour they are not going to use it."*

Steps 2B.1.2 and 2B.1.3 - Determine importance of criterion

The relative importance of criteria was considered by the formal students as well as by the professionals. The professionals briefly discussed the generation of criteria-importance weightings for individual criteria and then abandoned the idea. This was for simplicity reasons: *C: "But why wouldn't it be as quick to put a weighting on it as well?" B: "Just more maths ... just ... the numbers..." C: "... the time it's gonna take is determined by the weighting on it rather the calculation I'm thinking."* Later, the professionals thought about importance weightings again, but came to the conclusion that all criteria are equally important: *A: "Well, they are all important."* However, throughout the professionals' decision-making process it appeared as if the alternatives' evaluations restricted to particular criteria, especially 'efficiency of delivery', were paid more attention than other restricted evaluations.

In contrast to the professionals, the formal students went through steps 2B.1.2 and 2B.1.3 to establish formal criteria-importance weightings.

Step 2B.1.2, 'pick a criterion', was represented by the activity 'controlling the process'. For example: *C: "Okay, what about the importance of 'safety'?"* Step 2B.1.3 involved a variety of activity categories: 'weighting criteria formally', 'weighting criteria informally', 'raising evidence on criteria weights', 'defining criteria' and 'clarifying the product environment'.

The pattern of these activities when carried out as process step 2B.1.3, i.e. to determine the importance of a criterion, is shown in figure 4.16.

The activities as shown in figure 4.16 were carried out without any obvious *time* sequence. An exception is 'controlling the process' (actually step 2B.1.2) which most often initiated weighting a particular criterion. The activity 'weighting criteria formally' was always carried out when step 2B.1.3 occurred. The other activities were not carried out at all occurrences of this process step. Through studying the pattern of activities I perceived an apparent *functional* sequence: 'defining criteria' and 'clarifying the product environment' informed the activity 'raising evidence on criteria weights', which generated a justification for the outcome of the activity 'weighting criteria informally'. This outcome, i.e. an informally expressed weighting, was then mapped onto a formal scale by the activity 'weighting criteria formally'. Examples of transcribed activity patterns as modelled by figure 4.16 can be found in appendix E.

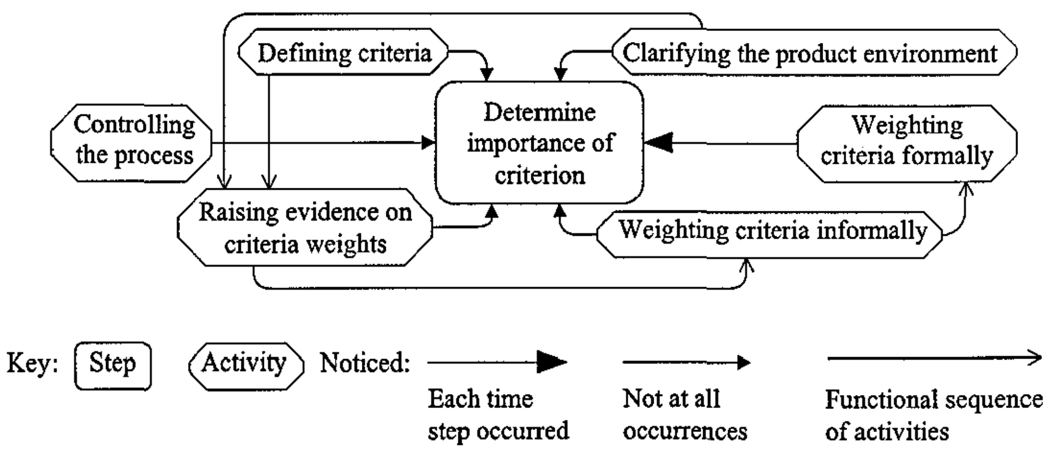


Figure 4.16: Determine importance of criterion (step 2B.1.3).

It was of particular interest to notice that not in all cases was evidence raised before a criterion was weighted (see figure 4.16). This implies that not all weights were justified. Yet, it seemed as if it was often implicitly understood by the group that a particular weighting at hand was justified by some utterance made earlier during the session. The group seemed to 'remember' evidence raised earlier. Altogether nine out of ten criteria weights appeared to be justified.

The activity 'weighting criteria formally' in many cases explicitly represented the mapping of informally expressed weightings onto a formal scale (see figure 4.16). The formal scale applied by the group consisted of the weighting factors '1', '2', '3', '4', and '5', where '5' represented utmost importance. These factors were seen as absolute ratings and treated as quantitative, cardinal data when the group used them to calculate comprehensive performances of alternatives by applying a weighted sum aggregation function (see table 4.1 and section 4.4.3).

The informally expressed weightings being mapped onto the formal scale were often relative rather than absolute (see figure 4.8) and qualitative or vaguely quantified, but never strictly quantitative (see figure 4.9). Hence, whenever such informally expressed weightings were mapped onto the formal scale a format transformation took place. An example for transforming relative weightings into absolute factors is: *C: 'I would go for 'running cost' and 'maintenance cost'...give them '3'." E: "Well, we said it is a lot more important than 'initial cost'* [which was weighted '2'] ... *'3' is not much difference to '2'." A: "Hmm, yeah I would go for '4' then."* An example for transforming vaguely, i.e. linguistically, quantified weightings into quantitative factors is: *C: "I think 'different cylinders' is quite important." E: "Yeah, make it five." D: No, not five." E: "Okay, make it four." A, C, D: "Yeah."*

Step 2B.2 - Evaluate alternatives restrictedly

The evaluation of alternatives restricted to criteria in step 2B.2 could be further broken down into three steps: 'pick an alternative', 'pick a criterion' and 'evaluate alternative-criterion pair'. These steps were sequenced by two different variants. In variant 2B.2A, which I called 'alternative-criteria', one *alternative* was picked and then evaluated with respect to all criteria before the next alternative was picked. That is, the picked alternative was evaluated independently from other alternatives. At variant 2B.2B, which I called 'criterion-alternatives' one *criterion* was picked and then used for the evaluation of all alternatives before the next criterion was picked. That is, the alternatives were compared to each other with respect to the picked criterion. These variants are shown in figure 4.17.

The formal students as well as the professionals went through step 2B.2 twice. The formal students first evaluated each alternative with respect to constraint satisfaction. For these evaluations variant A was applied. During the considerations of alternative-constraint pairs they examined whether or not the alternative at hand satisfied a particular constraint. The evaluation of such a pair was independent of other alternative-constraint pairs.

Subsequently, the group evaluated each alternative with respect to each objective. For these evaluations variant B was applied. The alternatives' performances were compared to each other, across all alternatives, and evaluated on the basis of a datum. Concept A was chosen as the datum.

The professionals first evaluated all alternatives with respect to each criterion, which resulted in the pre-selection of two alternatives. They then evaluated these two alternatives again with respect to each criterion. For the first evaluations, involving all alternatives they intended to apply variant A. This is because they had expressed the

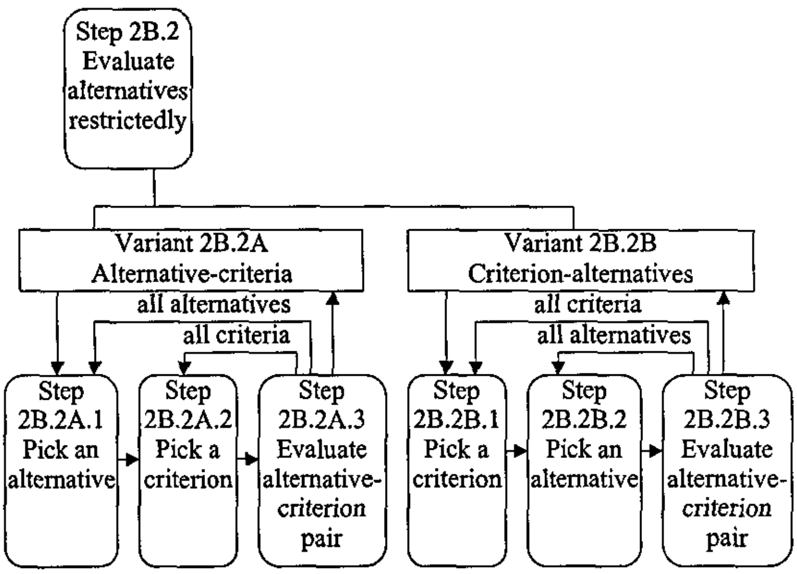


Figure 4.17: Evaluate alternatives restrictedly (step 2B.2)

intention to evaluate each alternative's performance relative only to a chosen datum's performance with respect to each criterion in turn. *C: "What we are comparing it against is the current method of delivery* [datum], *yeah?" A and B: "Yeah" C: "Right, we just gonna go through each one."* Yet, after the first alternative had been evaluated this way, the professionals developed a tendency to involve more alternatives and to compare performances across these alternatives rather than to the datum only. This means that, in fact, variant B rather than variant A was applied.

For the second evaluations, involving the two pre-selected alternatives only, the professionals expressed the intention to evaluate independently rather than through comparisons: *B: "We score each one 1 to 5 ... if you start using numbers ...you might get a clearer result come out ... it's just not gonna be compared with anything" A: "And what's the 1 to 5 gonna represent?" C: "1 is poor, 5 is excellent, yeah?" B: "Yeah!"* For these evaluations, variant B was chosen: *B: "Yeah, and maybe if we do ... as we go down each line* [for each criterion] *we do them together." C: "Yeah, okay." B: "So it's more honest..."* It seems that, for the generation of relative performance assessments, i.e. for the first evaluations, a variant was chosen that appears to be more suitable for the generation of absolute performance assessments (theoretically; without using a datum), whereas for the generation of absolute performance assessments, i.e. for the second evaluations, a variant was chosen that appears to be more suitable for the generation of relative performance assessments.

Steps 2B.2A/B.1, 2 and 3 - Evaluate alternative-criterion pair

In either variant, steps 2B.2A/B.1 and 2B.2A/B.2 were represented by the activity 'controlling the process'. Step 2B.2A/B.3 involved a variety of activity categories: 'gaining external information', 'determining the product environment', 'making assumptions on the product environment', 'discussing sub-issues', 'accepting assumptions on sub-issue solutions', 'clarifying concept working principles', 'defining criteria', 'identifying criteria', 'raising evidence on restricted performance of concepts', 'determining or evaluating restricted performances informally' and 'evaluating restricted performances formally'.

The pattern of these activities when carried out as process step 2B.2A/B.3, i.e. to evaluate an alternative-criterion pair, is shown in figure 4.18.

The activities as shown in figure 4.18 were carried out without any obvious *time* sequence. An exception is 'controlling the process' (actually step 2B.2A/B.1/2) which most often initiated evaluating an alternative-criterion pair. Through studying the pattern of activities I perceived an apparent *functional* sequence: 'gaining external information' aided 'determining the product environment'. If this environment could not be determined, due to lack of information, the groups were 'making assumptions on the product environment'. The activity 'discussing sub-issues' lead to 'accepting assumptions on sub-issue solutions'. The activities 'defining criteria', 'clarifying concept working principles', 'discussing sub-issues', 'accepting assumptions on sub-issue solutions', 'making assumptions on the product environment' and 'determining the product environment' all informed the activity 'raising evidence on restricted performance of concepts'. This activity, in turn, delivered a justification for the outcome of the activity 'determining or evaluating restricted performances informally'. This outcome, i.e. an informal evaluation, was then mapped onto a formal scale by the activity 'evaluating restricted performances formally'. 'Identifying criteria' was not carried out with any perceivable purpose during this process step. New criteria simply emerged as part of the evaluations. Examples of transcribed activity patterns as modelled by figure 4.18 can be found in appendix E.

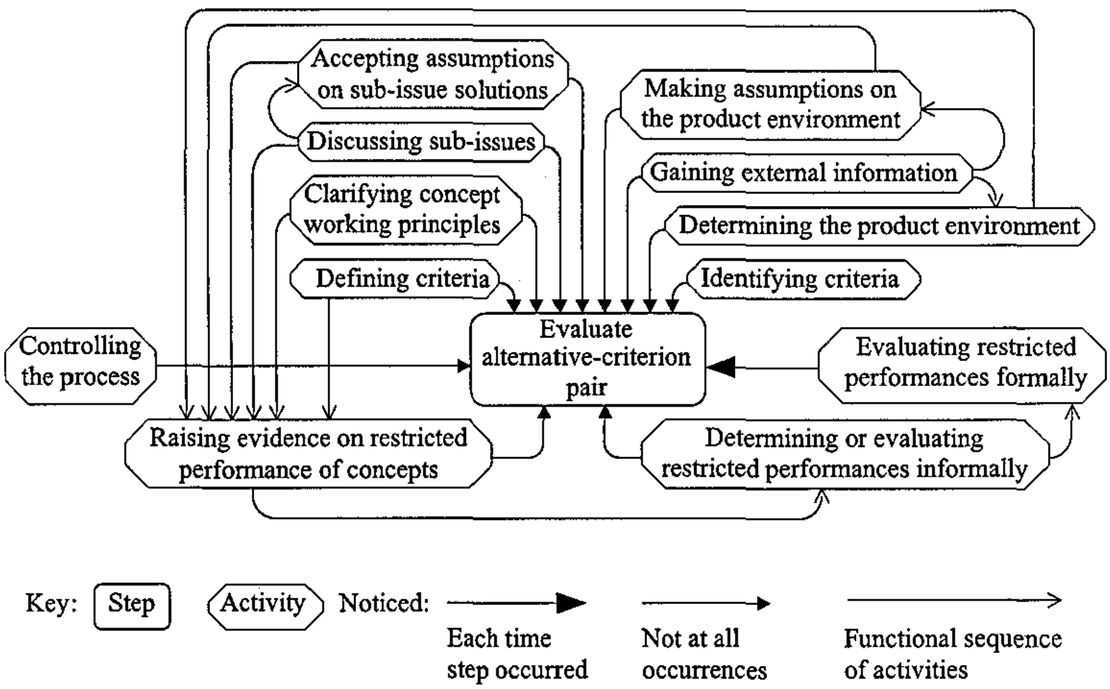


Figure 4.18: Evaluate alternative-criterion pair (step 2B.2A/B.3).

The step 'evaluate alternative-criterion pair' occasionally had an iterative nature. That is, restricted evaluations on particular alternative-criterion pairs that had been concluded were resumed at a later time and re-assessed.

The activity 'evaluating restricted performances formally' was always carried out when step 2B.2A/B.3 occurred. The other activities were not carried out during all occurrences of this process step (see figure 4.18). Figure 4.19 shows the relative number of this step's occurrences involving particular activity categories.

Figure 4.19 indicates that the prevailing activities for evaluating alternative-criterion pairs were 'evaluating restricted performances formally', 'determining or evaluating restricted performances informally', 'raising evidence on restricted performance of concepts' and 'controlling the process'.

Another activity that occurred frequently was 'discussing sub-issues'. Within discussions on sub-issues, alternative conceptual solutions for the sub-issue were brought forward and informally evaluated. When a possible solution was agreed upon, this was assumed to be part of the main concept and fed into the discussion on its performance with respect to the criterion at hand. Figure 4.19 indicates that 'accepting assumptions on sub-issue solutions' occurred less frequently than 'discussions on sub-issues'. This implies that a number of these discussions did not have any explicit outcome.

The two activities 'defining criteria' and 'identifying criteria' were involved in evaluating alternative-criterion pairs only in the professionals' decision-making process (see figure 4.19). 'Defining criteria' aimed at clarifying the definition of the criteria's meaning and reminding individual group members about definitions established earlier. Especially the criterion 'ease of use - weight to be handled' appeared to cause some difficulties regarding its definition. Discussions on its meaning were raised a number of times. Although the group seemed to have agreed on a particular definition during process step 2B.1.1, they started discussing it again during the evaluations of alternative-criterion pairs during process step 2B.2A/B.3.

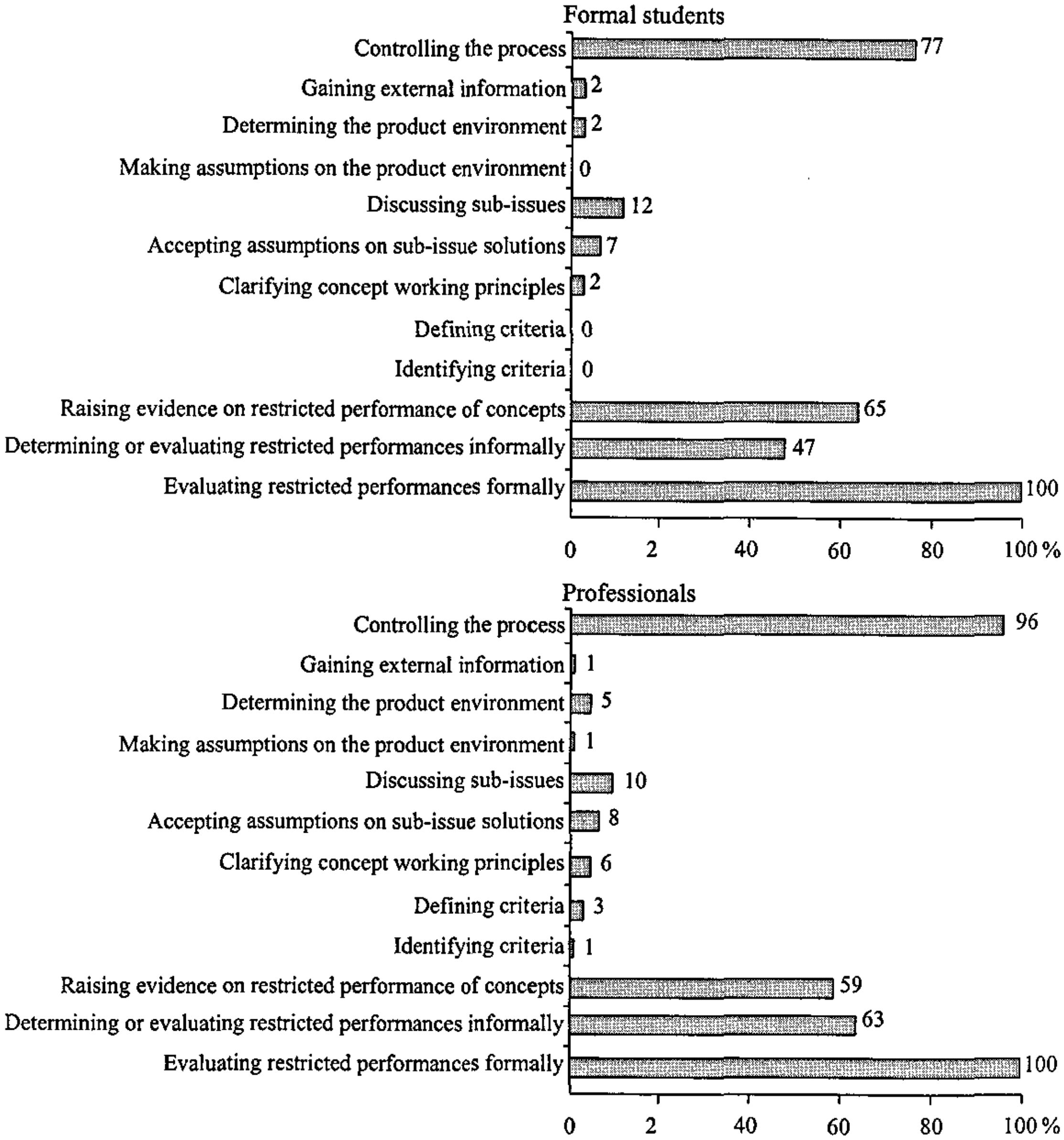


Figure 4.19: Relative occurrence of step 2B.2A/B.3 involving particular activity categories.

During the evaluation of alternative-criterion pairs, the professionals also identified the new criterion 'ease of development', which they considered as being very relevant. Yet, this criterion was not added to the group's criteria structure and as it was identified late in the process it was decided to simply ignore it for the current decision-making process: *C: "I think we stick* [to the current set of criteria]... *we have got a time constraint* [for the decision-making process]. *It* [the new criterion] *doesn't have to be part of that initial assessment."*

As shown by the model in figure 4.18 and the percentages in figure 4.19, evaluating alternative-criterion pairs did often, but not always involve the activity 'raising evidence on restricted performance of concepts'. A lack of this activity implies that some of the formal evaluations were made without any justification. Yet, sometimes evidence had actually been raised, but not made explicit during the evaluation of a particular alternative-criterion pair. This is because (i) discussing and accepting assumptions on sub-issues, clarifying the concept working principles as well as determining and making assumptions on the product environment sometimes led to implicit justifications for performance assessments and evaluations. Occasionally, apart from this, evidence raised for the justification of one alternative-criterion pair was also valid for evaluations of other alternative-criterion pairs. For example, the evidence: *A: "...telescope* [concept A's telescope rail mechanism] *is rather complicated..."* may by applied for justifying performance assessments restricted to 'initial cost', 'running cost', 'weight to be handled', 'time of each delivery' and 'training required'. Such evidence was not explicitly mentioned again subsequent to its first mentioning, as the group seemed to remember it.

Figure 4.19 shows that the activity 'determining or evaluating restricted performances informally' occurred very frequently during the formal students' as well as during the professionals' evaluations of alternative-criterion pairs. The outcome of this activity is an informal evaluation. As these two groups used formal evaluation matrices, any informal evaluations had to be mapped onto the groups' pre-defined formal scales. This mapping was represented by the activity 'evaluating restricted performances formally'.

The formal students first evaluated all alternatives with respect to constraint satisfaction. Constraint satisfaction was assessed on a binary basis. These were independent, qualitative statements saying whether or not a particular constraint was considered satisfied by a specific alternative. During these evaluations there were no informal expressions of a format different from the required independent, i.e. absolute rather than relative, qualitative binary statements. Hence, there was no mapping of informal evaluations onto formal scales.

Once the alternatives were assessed with respect to constraint satisfaction, they were evaluated with respect to each objective in turn. For these evaluations a formal scale consisting of the scores: '-2', '-1', '0', '+1' and '+2' was chosen. As the comparison basis one alternative was selected as the datum upon which all other alternatives' performances were assessed. A negative score indicated a performance worse than the datum's whereas a positive score indicated a performance better than the datum's.

Because these scores refer to performance comparisons they are of a relative nature. Apart from this, the scores were qualitative, e.g. '-2' did neither mean two times worse than '-1' nor any other quantification of a performance difference; it was simply interpreted as 'worse than '-1''. During these evaluations the activity 'determining or evaluating restricted performances informally' occurred frequently. This means that there were evaluation statements that had formats different from qualitative-relative. These were, in particular, vague quantifications (see figure 4.9) and statements indicating likelihoods (see figure 4.10). Any of such statements had to be transformed onto the formal qualitative scale. Yet, information on likelihoods could not be transformed at all.

An example for the transformation of vague quantifications is: *E: "This* [concept B] *is not anywhere near as expensive as that* [datum], *but compared to C* [concept C] *it's more expensive." D: "Okay, '+1'* [for concept B]*."* In this example, the vague linguistic quantification 'not anywhere near as expensive' was transformed into the quality 'less expensive' or 'better', i.e. '+1'. All quantitative content, even though being vague, was lost during the transformation.

The professionals also applied the datum method. This was for their first evaluations involving all alternatives. Initially they intended to use a scale consisting only of the scores '-1', '0' and '+1'. Yet, during the evaluations, when they developed the tendency to compare across all alternatives rather than with respect to the datum only, this scale was refined by the addition of the score '-2'. The same transformations occurred as during the formal students' evaluations regarding the alternatives' performances with respect to the objectives.

For evaluating the pre-selected two alternatives the professionals applied a formal scale consisting of the scores '1', '2', '3', '4' and '5'. These scores did not intend to represent *relative* differences between alternatives, but independent, absolute evaluations. A score of '1' was interpreted as 'poor' and score of '5' was interpreted as 'excellent'. Scores *between* '1' and '5' were interpreted as evaluations *between* 'poor' and 'excellent'. Although a linguistic definition, such as 'excellent', inherits vagueness, the vagueness was not actually modelled. The scale creates five sets: 'poor', ... , 'excellent', which are *labelled* with vague linguistic terms. However, the sets themselves are crisp because they only allow any evaluation, mapped onto them, to either be a member or not. This mapping means assigning the evaluation an associated set. Once an evaluation has been assigned to one of these sets it will be given a crisp numeric value, e.g. '5', which is the same for all members of this set. Therefore, the format of these restricted, formal evaluations is quantitative apart from being absolute.

There were frequent occurrences of the activity 'determining of evaluating restricted performances informally'. This means, there were evaluation statements that had formats different from quantitative, absolute. These were qualitative or vague quantifications, rather than strictly quantitative, or relative rather than absolute or they indicated likelihoods. Any of such statements had to be transformed onto the formal quantiative, absolute scale.

An example for the transformation of a relative, vague quantification indicating likelihood is: *B: "Em, I'd say C* [concept C] *is a little bit better* [than concept B]*." A: "Definitely better...." C: "So it's ... two and three* [scores]*."* In this example, the two alternatives are actually compared to each other rather than assessed independently. So, it becomes clear why they were assigned different scores, but it is not clear why it was '2' and '3'. The vague quantification 'a little bit better' was transformed into the precise quantity '1' (difference between concepts C and B). This pretends a precision that was not actually inherent in the original vague quantification. The likelihood expressed through 'definitely better' could not be transformed at all.

4.4.3 Conclude the process

In any case, whether the alternatives were evaluated using the comprehensive approach in step 2A or the criteria-based approach in step 2B, all groups concluded their decision-making processes by first establishing comprehensive ranking orders and then analysing these orders before their final decisions were announced. This is process step 3, which could be disintegrated into Step 3.1, establish ranking, and step 3.2, analyse ranking. These two steps are shown in figure 4.20.

Step 3.1 - Establish ranking

The informal students had not applied any formal criteria-based evaluations in step 2. Instead, comprehensive evaluations were directly made. This means that no further procedures were needed to aggregate restricted, criteria-based evaluations into comprehensive evaluations.

They established an overall ranking order by initially classifying the alternatives under three different labels: the two best alternatives, the worst alternative, and the alternative that is somewhere in-between. No evidence was explicitly raised that could have justified the relative comprehensive evaluations leading to this ranking order. The two best alternatives were then ranked again relatively to each other on the basis of their performance with respect to one particular criterion. Evidence was raised that justified the superiority of the one alternative over the other. Raising this type of evidence was represented by the activity 'raising evidence on comprehensive performance of concepts'. Thus, a complete ranking order was achieved. The order was expressed by assigning each alternative an ordinal ranking number, i.e. an integer

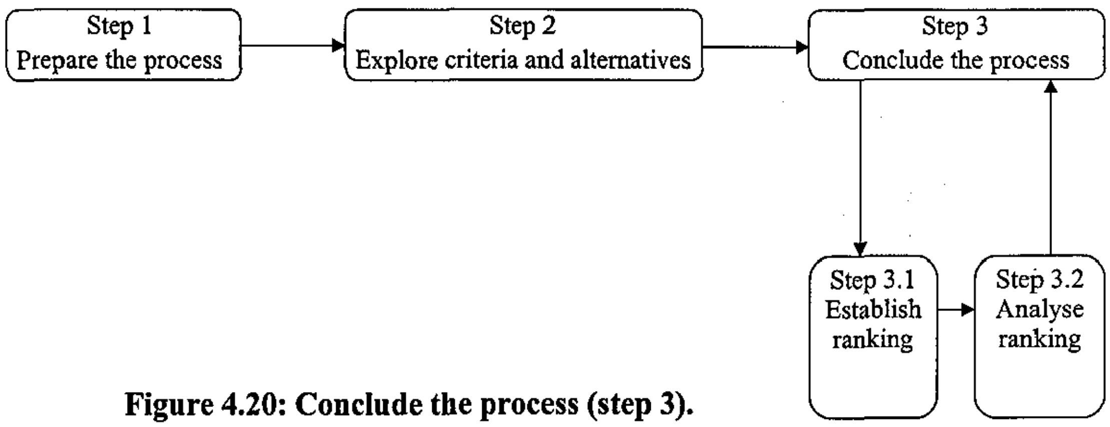


Figure 4.20: Conclude the process (step 3).

in the range from 1 to 4 meaning best to worst, respectively. This assignment of ranking numbers was represented by the activity 'evaluating comprehensive performances formally'.

The formal students had evaluated the alternatives using the criteria-based approach in step 2. This means that they first had to aggregate the restricted criteria-based evaluations into comprehensive evaluations before an overall ranking could be generated. This aggregation was achieved by calculating an overall, comprehensive, score for each alternative. To do so a weighted sum aggregation function was applied, i.e. the scores for each alternative-criterion pair were multiplied with the associated criteria-importance weights and then summed up for each alternative. The higher the sum, the better was the comprehensive performance of an alternative. These sums were then transformed into a ranking order. Calculating comprehensive evaluations was represented by the activity 'evaluating comprehensive performances formally'.

The scores applied for the restricted criteria-based evaluations were of a qualitative, relative format (see section 4.4.2). This means that the formal students' calculations using a numeric weighted sum aggregation function were based on qualitative, relative scores, i.e. on ordinal data! The aggregation of restricted criteria-based evaluations into comprehensive evaluations can be seen in table 4.1.

The professionals used a similar procedure as applied by the formal students for their first evaluations, which lead to the pre-selection of two alternatives. However, the professionals aggregated the restricted criteria-based evaluations into comprehensive evaluations by using a simple sum aggregation function, rather than a weighted sum, i.e. no criteria-importance weights were involved. This can be seen in table 4.2.

For the comprehensive evaluations of the pre-selected two alternatives the professionals again used a simple sum aggregation function. This time the calculations were based on quantitative, absolute scores which is a valid calculation base and may even suggest an overall rating rather than a ranking. This can be seen in table 4.3. However, these quantitative, absolute scores are questionable as they had largely been achieved through transforming qualitative or vague quantitative, relative evaluations (see section 4.4.3).

Step 3.2 - Analyse ranking

Once the ranking orders were established the three groups proceeded by analysing them in step 3.2 (see figure 4.20). This analysis either concentrated on deliberating various sub-issues addressing the highest ranked alternative or on mapping group members' individual, intuitive rankings onto the one established by the entire group. The former type of analysis, which was adopted by the informal students, seemed to have the purpose of thoroughly checking the highest ranked alternative's feasibility. The latter type of analysis, which was adopted by the formal students and the professionals, seemed to have the purpose of checking whether all individual group members were satisfied with the calculated rankings. Hence, it was an attempt to understand why particular alternatives had achieved a specific position in the calculated ranking order.

The formal students' and the professionals' analyses were based on the activity 'mapping intuition onto ranking'. That is, the individual decision-makers mapped their intuitive rankings onto the calculated ranking and, if a mismatch was perceived, the group attempted to clarify the cause. Through the activity 'raising evidence on comprehensive performance of concepts' evidence was gathered that could justify the relation of a specific alternative to other alternatives within the comprehensive ranking order. This type of evidence was in many cases represented by restricted criteria-based evaluations, first established in step 2B.2A/B.3, and their underlying justifications. For gaining such evidence the two groups very frequently referred to their completed evaluation matrices. Yet, because these matrices only contained formal scores, their underlying justifications had to be recalled or reconstructed.

If mismatches between intuition and calculation could be resolved and the group members who perceived the mismatch were convinced about the calculated ranking's correctness, the highest ranked alternative could be selected. If the evidence was not convincing, another restricted evaluation process was suggested. This happened to the professionals who doubted that the alternative ranked highest in the first evaluations really was the best alternative and hence, saw indeed a need for repeating some evaluations. Figure 4.21 shows the formal students' and the professionals' analysis of their ranking orders.

At the end of each workshop I asked the groups if they considered the selected alternative as generally effective or simply as the best among the available alternatives. This question aimed at finding out whether they would recommend either searching for further alternative solutions or detailing the selected alternative.

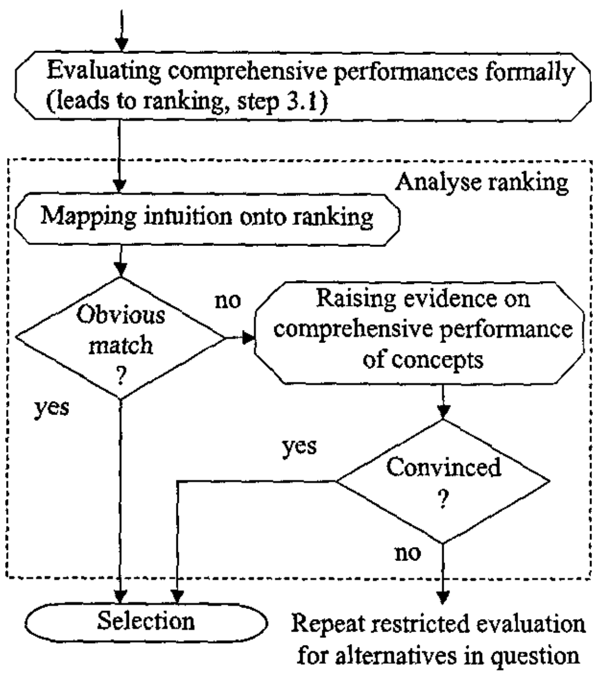


Figure 4.21: Analyse ranking (step 3.2, formal students and professionals).

The informal students claimed, as an immediate reaction, that they had evaluated the alternatives on an absolute basis and that therefore their selection can be considered sufficiently effective. Yet, after some more deliberations on this aspect the group eventually withdrew their earlier claim with the statement *B: "...well, from our concepts, we would definitely say concept A was the best...."* The formal students suggested that their selection represented the best out of the pool of alternatives as given to them. For further development they recommended to combine two of the evaluated alternatives. The professionals stated that their selection was not considered satisfactory on an absolute basis. The search for further, more effective alternatives was recommended. Both the formal students as well as the professionals suggested to further develop the alternatives with respect to particular sub-issues.

A model summarising the decision-making process as discussed throughout section 4.4 is shown in figure 4.22.

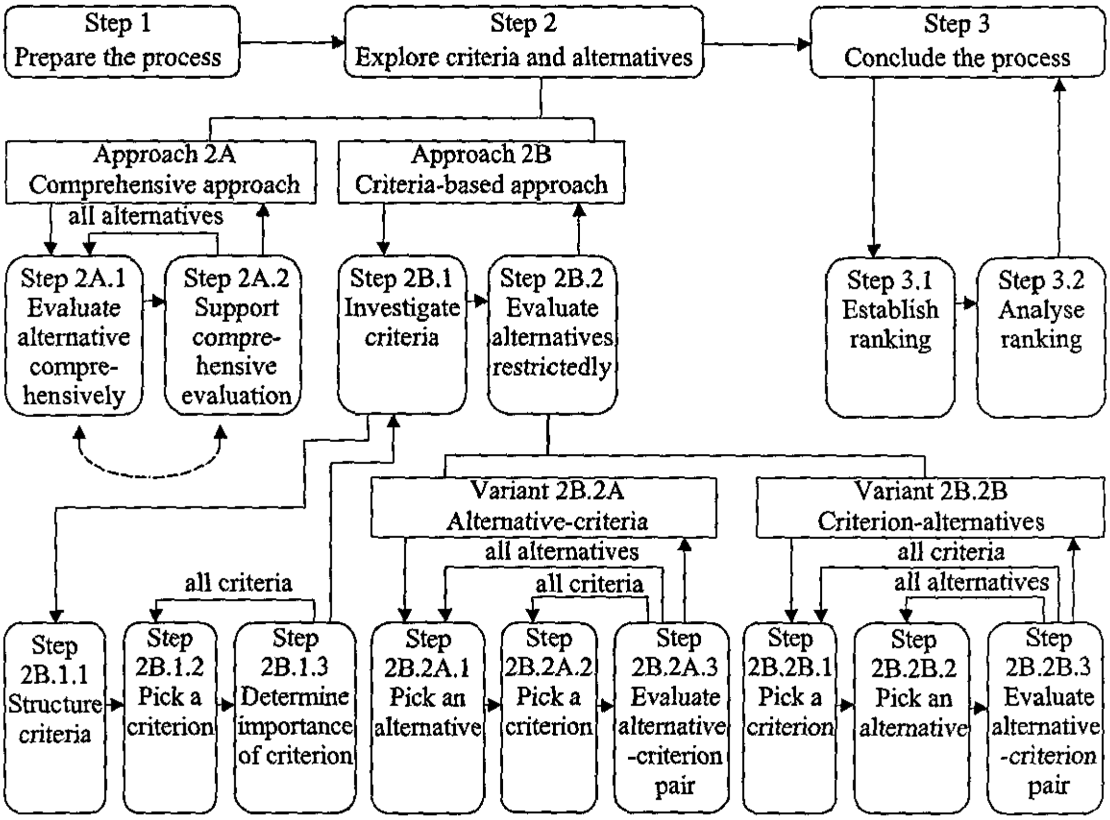


Figure 4.22: Decision-making process model.

4.5 Summary

The aim of this chapter was to introduce the research results. Basically, these were a set of decision-making related activity categories and a decision-making process model.

The set of activity categories consists of twelve main categories and fifteen sub-categories. Confidence in this set was gained by stabilising it across the three analysed transcripts and by discussing its accuracy.

The discussion on accuracy was based on a number of previous studies on design processes. It was found that many of my categories could be associated with categories or observations from the previous studies. This provided confidence. Some of my categories could not be exactly associated or not be associated at all with categories or observations from the previous studies. This means that I could extend the set of known activity categories related to design processes, with the specific focus on decision-making in design.

Through investigating each activity category's time consumption across the three groups and through studying occurring formats of particular information, specific process characteristics were identified.

More process characteristics were then identified through generating a detailed decision-making process model comprising all three groups' approaches and different process variants. This model indicates the general decision-making process structure as well as very specific relationships between different activity categories within particular process steps.

Within this chapter only the research results, but not their implications have been addressed. A discussion of implications will be the focus of the next chapter.

Chapter 5

Implications of Results

5.1 Introduction

This chapter will discuss the implications of the study's results. These are expressed as observations and finally as a set of requirements for decision-making support. All observations generally address the research objective 2.1, as first stated in section 1.2. Some observations also specifically address the guiding research questions, related to objective 2.1, as stated in section 2.4.4. The set of requirements addresses the research objectives 2.2 and 3.1, as stated in section 1.2. Most sections, figures or tables referenced within the following discussions refer back to chapter 4, 'Results and Analysis' or chapter 2, 'Review of the Literature'.

5.2 Observations

This section will introduce and discuss altogether twelve individual observations that I made when analysing the three groups' decision-making processes. These observations address particular process characteristics across the three groups. Many observations were triggered by the guiding research questions, raised in section 2.4.4. In return, the observations, as discussed below, will deliver answers to all of these questions.

Observation 1: core activities

In section 4.3.3, it was shown that for each group there were six activity categories that together consumed 90% of the groups' entire process time. I called them Pareto 'A+B elements'. Interestingly, five of these six categories were identical across the three groups. These categories were:

- Discussing the process approach;
- Deliberating sub-issues;
- Raising evidence;
- Determining or evaluating performances;
- Controlling the process.

These five categories seem to represent core activities in the observed decision-making processes, regardless of the chosen approach and the group members' professional experience.

The observation confirms findings of Dwarakanath and Wallace (1995) / Dwarakanath (1996) with respect to *groups*. They observed individual designers and groups while generating conceptual designs. Dwarakanath and Wallace found that considerable time was spent on 'arguments process' (my categories 'discussing the process approach' and 'controlling the process') and that 'deliberating sub-issues' plays an important role. For the *individual* designers they measured very little time consumption for 'arguments product' (my categories 'determining or evaluating performances' and 'raising evidence'). However, with respect to design *groups* Dwarakanath (1996) stated that "alternatives tend to be continuously evaluated..." (p.79). This is reflected in my groups by their core activities directly related to the evaluation of alternatives, i.e. 'determining or evaluating performances' and 'raising evidence'. This, in relation to Dwarakanth and Wallace's observation, seems to suggest that groups dedicate more time for explicit evaluation activities than individuals.

Observation 2: defining criteria

Apart from sharing the above five core activities, the formal students and the professionals also share the sixth category in their sets of 'A+B elements' (see section 4.3.3). This category is 'defining criteria'. However, 'defining criteria' did not have much time consumption for the informal students. Instead, the sixth category in their set of 'A+B elements' is 'gaining external information', which, in turn, did not have much time consumption for either the formal students or the professionals.

Feldy (1997), speaking from years of practical design experience, argues that efforts to address criteria (e.g. 'defining criteria') are a sign of effective decision-making because it helps to really understand the actual goals of the process. Criteria and goals describe the decision problem rather than its solutions. Jebb and Woolliams (2000, p.281) phrase it this way: effective processes in design involve the exploration of *"problem land before solution land"*. Therefore, my observation that 'defining criteria' was a core activity for the two groups who applied formal decision-making approaches, but not for the informal group, suggests that using such approaches may positively influence the groups' effectiveness.

Observation 3: experience and approach

Even though there is agreement with respect to five core activity categories, being in all sets of 'A+B elements', the individual time consumptions of categories are notably different across the three groups' processes (see figure 4.4). Considering the relative times spent by each of the three groups on the activities classified by the categories in their sets of 'A+B elements', I observed that the formal students and the professionals had more similarities to each other than to the informal students (see also the visual 'pattern' of the three charts in figure 4.4). This was despite the different levels of professional experience of these two groups, i.e. one group being students and the other group being professional engineers.

Regarding the sub-categories' time consumptions, the formal students also had more similarity to the professionals than to the informal students (see figures 4.5, 4.6 and 4.7).

The observation suggests that the effects of using a formal approach, as by the formal students and the professionals, might override possible effects of different professional experience.

Observation 4: raising evidence

Having observed that the formal students had more similarities with the professionals than with the informal students, there was an interesting exception. That is, with respect to 'raising evidence', the formal students seemed to have more similarity to the informal students rather than to the professionals. The professionals spent considerably more time on 'raising evidence' than either of the student groups (see figure 4.4).

Students and professional engineers have also been compared in an observational study by Smith and Leong (1998). They found that professional engineers paid more attention to details than the student groups. This may have been the underlying reason for my observation that the professionals spent considerably more time on raising evidence than the students.

Further findings of Smith and Leong (1998) could not be confirmed by my study. They found that, throughout the process, the professional engineers engaged in more management activities than the students. I did not notice such behaviour. Management activities were categorised in my study as 'discussing the process approach' and 'controlling the process'. Figure 4.4 shows that the professionals did not spend more time on these activities than the students, at least not more than the formal students. In Smith and Leong's (1998) study no formal methods were apparent. Therefore, I suggest that it may be the use of a formal approach that resulted in extensive management activities in both of my formal groups.

Observation 5: criteria importance

The groups spent relatively little or no time at all on weighting criteria and raising evidence on criteria weights. This implies that little time was spent on the overall aspect of criteria importance.

This observation is particularly interesting in the light of Ehrlenspiel and Dylla's (1993) finding that 'successful' designers thoroughly analyse demands, which includes defining their importance. I found that none of my groups naturally showed this quality.

The only group that explicitly considered different levels of importance for the criteria was the *formal* students. This seems to suggest again that the application of formal approaches had a positive influence on the groups' decision-making process.

Nevertheless, considering the aspect of criteria importance appeared to be little more than a formality. This is because surprisingly little time was dedicated to the activities related to this aspect, i.e. 'weighting criteria' and 'raising evidence on criteria weights'.

Table 4.5 shows the criteria addressed by the informal students. The table indicates that particular criteria were treated in different ways: Whereas for example 'efficiency of delivery', was repeatedly applied for the evaluations of different alternatives, the criterion 'cost' was not considered any more after the first alternative's evaluation. This may suggest that the group perceived the one criterion as more important than the other one. Yet, this was never made explicit.

Observation 6: consistency

During the informal students' process I observed a comprehensive approach for exploring the criteria and alternatives. This approach did not involve much exploration of criteria (see figures 4.12 and 4.22): they would pick an alternative and comprehensively evaluate it. They also gathered supporting advantages and disadvantages, but they never used the full set of evaluation criteria. Moreover, as can be seen in table 4.5, the evaluations of different alternatives were not based on the same criteria, which suggests bias. Table 4.5 also shows that the gathered advantages and disadvantages sometimes referred to high level criteria and at other times referred to more specific low-level criteria. Thus, specific factors, represented by the low-level criteria, which had an influence on the alternatives' performance with respect to the higher level criteria, were ignored.

It was interesting to notice that the informal students dedicated more time to evaluating the alternative that was first in the list than for any other alternative. This may have influenced their decision because they finally selected this first alternative as their preferred solution.

Ullman et al. (1988) as well as Dwarakanath and Wallace (1995) also empirically studied informal designers. Ullman et al. (1988) found that designers usually focus on only a few criteria to reduce the evaluation's complexity and that criteria were occasionally forgotten and therefore not addressed during the alternatives' evaluations. Dwarakanath and Wallace (1995) noticed similar behaviour. These two studies therefore offer support for my observation.

The formal criteria-based approaches of the formal students and the professionals ensured that all criteria were consistently considered for the evaluation of all alternatives. This consistency may have removed the bias addressed above and may be the explanation for both formal groups having selected the same alternative (which was not the first alternative), independently from each other.

Ehrlenspiel and Dylla (1993) claimed that a characteristic of less successful designers is being less concrete and rather emotional when assessing solution properties. I observed a higher level of rigour for those groups that applied formal approaches. This rigour can be seen by their consistency in including of *all* criteria in the evaluation of concepts. This was in contrast to the group that used the informal, comprehensive approach. Therefore, those groups that used formal evaluation approaches apparently showed characteristics of successful designers, whereas the group that did not use any formal methods apparently showed characteristics of less successful designers.

Observation 7: absolute vs relative evaluations

In the formal, criteria-based approaches there were two variants of evaluating the alternatives (step 2B.2): 'variant A', alternative-criteria and 'variant B', criterion-alternatives (see section 4.4.2, figure 4.17). Basically, variant A leads to independent, absolute evaluations and variant B leads to evaluations by comparison.

I observed that absolute evaluations were only successfully carried out if the evaluations were based on binary assessments. This means, an alternative may satisfy or not satisfy a criterion. This was done when the formal students checked the alternatives for constraint satisfaction or when the informal students comprehensively evaluated the alternatives as being 'good' or 'not good' backed by advantages or disadvantages. However, whenever the groups tried to express degrees of satisfaction that were not binary, they did this by comparing alternatives (see section 4.4.2, step 2B.2A/B.3).

The professionals attempted to express degrees of satisfaction based on independent, absolute evaluations. However, this attempt was unsuccessful and ended up in comparisons again. The reason seemed to be that independent, absolute evaluations require some sort of effectiveness function (see section 2.2.1) to 'measure' the alternative performances' effectiveness. Such functions did not appear to exist. In other words, the group did not exactly know what performance may be independently rated as '4' (very good) or '5' (excellent). But they did know whether one alternative's performance was better or worse than another alternative's performance. Usually, these comparisons were expressed qualitatively or in vague quantiative terms, but they were hardly ever precisely quantified (see section 4.3.4, figure 4.9).

Simply comparing alternatives with respect to criteria appears to be straightforward and does not seem to cause problems for the decision-makers. This has also been found in other studies (Ehrlenspiel & Lenk, 1993; Weiss & Hari, 1997). Yet, such comparisons only relate the alternatives to each other rather than to an overall goal. This means that it can be established which alternative is the best. But the question on whether this alternative is sufficiently effective and should be further detailed or whether new, more effective alternatives should be produced cannot be directly answered. According to Feldy (1997) decision-making processes often fail in this respect. That is, when it comes to judging the overall outcome.

When I asked the groups whether they would recommend developing more alternatives, or going ahead with the selection, they could answer the question generally, but they could not exactly point out specific aspects that were unsatisfactory about the available alternatives, on an independent scale. The only question that they could answer, based on the information produced by their decision-making methods, was: 'which is the best alternative?' Ullman et al. (1997) claimed that decision-making methods traditionally only addressed this question rather than giving guidance for any further considerations.

In their empirical studies Ullman et al. (1988) observed that when designers compared different alternatives, one was taken as a datum, and the others were only compared to the datum rather than to each other. I did not observe this behaviour in my study. The informal group considered the alternatives individually and never explicitly used a datum. The formal groups did evaluate by comparing performances, but across a number of alternatives. The use of a datum as the only comparison basis did not seem practical: the professionals intended and started to evaluate their alternatives by exclusively comparing each one to a datum but, they gradually involved more alternatives in these comparisons. I see the formal groups' tendency to compare performances across all alternatives as a result of applying an evaluation matrix, which clearly laid out all alternatives in front of them. In contrast, the participants in the study conducted by Ullman et al. (1988) did not use a formal evaluation method.

Observation 8: justification of evaluations

The formal students' and the professionals' evaluations took place in the process step 'evaluate alternative-criterion pair' (see figures 4.17 and 4.18). Justifying evidence for the evaluation of an alternative-criterion pair was produced by the activity 'raising evidence on restricted performance of concepts'. I observed that this activity was not carried out consistently for all alternative-criterion pairs by either group (see figure 4.18). This resulted in a number of evaluations being seemingly unjustified; discussions had been made in an intuitive manner, which makes it difficult to retrospectively understand and justify them.

All groups were given the set of alternative concepts presented as sketches drawn on paper. The formal groups chose to use additional documentation by drawing evaluation matrices, clearly listing the criteria. It was interesting to notice that the *informal* group was only consistent with respect to addressing all alternatives. The *formal* groups were only consistent with respect to addressing all alternatives and with respect to applying all evaluation criteria. The implication of this is that the groups were only consistent with respect to those aspects of the decision-making process that were drawn up or documented. With respect to all other aspects, such as delivering justifying evidence, the groups were inconsistent. Such inconsistency can be interpreted as lack of accuracy and having a tendency towards being superficial. According to Ehrlenspiel and Dylla (1993) these are clearly characteristics of less successful designers.

Feldy (1997) believes that a benefit of using decision-making methods is the provision of defensible reasons for a decision. I agree, but this could only be partly observed in my study despite some of the groups having used a method. I suggest it is not the use of a method *per se*, but the need for explicit documentation that ensures the consistent provision of defensible reasons for a decision.

Observation 9: information capture

During all three groups' decision-making processes considerable amounts of valuable information were gathered. Basically, all activities shown in figure 4.4, apart from 'controlling the process', generated information. I believe this information had the potential to be re-used in later design activities. However, as most of it was not captured by the groups in any form, it seems unlikely that this information would have been available for later re-use.

The formal students as well as the professionals produced some records by completing their evaluation matrices. However, these matrices only contained formal evaluation statements (scores), but no underlying justifications, assumptions, deliberated sub-issues, etc. The group that did not use any formal evaluation method did not produce any records at all.

Ehrlenspiel and Lenk (1993) made a similar observation in their studies. Court (1998) reported on key findings from a number of empirical studies. Among these findings is that there was a clear lack of formal records in design processes. Furthermore, Shah (1998) claims this is a shortcoming observable in many companies.

Schlüter (1999b), speaking from many years as head of different design departments in industrial companies, believes the problem with producing records is simply that it takes time. In fact, one of the professionals suggested that they should take some notes; in particular on assumptions made (see activity 'making assumptions on the product environment'). However, nobody made the effort to actually do it. Ullman (2000) thinks that engineers spend a great percentage of their time recreating prior work or looking for prior information. He believes that to capture the full range of design information and to archive and query would be of extensive benefit for design efficiency and quality.

Observation 10: information formats

The formal students as well as the professionals partly generated information in formats that could not be directly accepted by their formal criteria weighting and performance evaluation scales. Therefore, such formats had to be transformed, which distorted their original meaning. These were, in particular, linguistically vague quantifications and expressions of confidence (see section 4.4.2).

Linguistically vague quantifications were statements that included linguistic quantifiers, which were used in a wide range. Ehrlenspiel and Lenk (1993) described experiments in which designers of different backgrounds (students and professionals) informally evaluated design solutions. They also noted that evaluations are usually expressed in vague, imprecise terms as in my study.

For my groups, confidence in evaluations meant how likely it was that an evaluation was actually true. They never expressed such likelihood numerically in percent, but always by using linguistic terms, which were most often associated with a (subjective) likelihood of 100%. Ullman et al. (1997) also described the occurrence of confidence expressions. They described expressions conveying *various degrees* of confidence. This was different in my study because I could only observe the exceptional expression conveying a confidence 'less than 100%'. But there is a difference between the meaning of 'confidence' in the studies of Ullman et al. (1997) and mine. In their study 'confidence' meant how likely it was that an absolute evaluation was true, whereas in my study 'confidence' refers to as how likely it is that a comparative evaluation is true.

Observation 11: bias vs time

As mentioned above, I noticed that the informal group spent a much higher proportion of time on evaluating the alternative that was first in the list than on any other alternative. This first alternative was eventually also selected by them. In contrast, the two formal groups spent similar proportions of time on the evaluation of each alternative and eventually selected, independently of each other, the same alternative. This selected alternative was not the first in the list (see tables 4.1, 4.3). It was also different from the informal students' selection.

Weiss and Hari (1997) argued that the selection of design concepts is problematic in most practical cases which is why they recommended the use of methods. They claimed that industrial practitioners commonly select the first proposed concept without thoroughly checking the alternatives. This behaviour imposes the risk of selecting a poor concept, which may have substantial consequences up to total failure of the entire product development program (Weiss & Hari, 1997).

Using formal methods as opposed to not using them seems to help in reducing bias, but it implies considerably longer decision-making processes, which was evident in my study (see section 4.2). Schlüter (1999a), arguing from an industrial practitioner's perspective, believes that the time taken to use methods must be reduced as it slows down the overall design process. Moreover, he sees the methodical evaluation of different alternatives as usually unnecessary because the best alternative is apparent (Schlüter 1999b). However, as my three groups did not all select the same alternative there was obviously no apparent 'best' alternative in my study.

Ehrlenspiel and Dylla (1993) observed that successful designers take their time when analysing and evaluating alternative solutions: they apply adequate and meaningful strategies, whereas less successful designers are less accurate and analyse on a superficial level. Feldy (1997) claimed that, despite their time consumption, formal decision-making processes are beneficial: they force the consideration of multiple alternatives early in the design process and build commitment for decisions.

Observation 12: iterations

I observed that all 3 groups occasionally repeated the evaluation of concepts that had been evaluated before (see section 4.4.2, steps 2A.2, 2B.2A/B.3). Such iterations happened regardless of whether the criteria-based approach or the comprehensive approach was used.

Vetschera (1994) questioned that decision-makers have a consistent system of preferences at the outset of decision-making processes. Instead, these seem to 'evolve' during this process, which may render iterations inevitable.

5.3 Requirements for decision-making support

The study's results suggest a number of requirements for decision-making support. I could identify five major aspects regarding such requirements:

- Building a process framework;
- Establishing a criteria structure;
- Handling the alternatives;
- Describing the product environment;
- Evaluating the alternatives.

For each stated requirement I will discuss (i) if and how it is addressed by the methods as introduced in chapter 2 and (ii) if it can be associated to any requirements mentioned or addressed in the identified previous studies in this field, reported by Dwarakanath (1996) and Ullman and D'Ambrosio (1995) (see section 2.4.3 and appendix B for details).

5.3.1 Building a process framework

There were, in particular, two general factors that lead to the very distinct approaches of the informal students on one hand and of the formal students as well as the professionals on the other hand. These factors were structure and information capture.

Structure

The two formal groups' approaches included the definition of an underlying process framework that provided structure and set an agenda. This agenda guided the decision-makers through their decision-making processes.

Observation 11 suggests that the apparent advantage of not applying a formal approach is time savings during the decision-making process. The informal students' process time of 45 minutes was much shorter than the two formal groups' (formal students and professionals) process times. They took 60 minutes and 80 minutes to resolve exactly the same task. However, I see the informal students' process as being flawed, because the different alternatives were not evaluated on the basis of the same set of criteria (see observation 6), which suggests 'unfair' evaluations and bias.

The aspect of bias renders the informal approach questionable. I doubt that the advantage of saving time outweighs the disadvantage of introducing bias. This is because it seemed to have influenced the final decision: observation 6 states, that the informal students selected a different alternative than the two formal groups, who both selected exactly the same alternative, based on 'fair' evaluations. Moreover, the informal students did not even consider the formal groups' selection as second best and vice versa.

Requirement: Decision-making support should encourage applying formal processes that provide an effective structure and set a guiding agenda. Benefit: The agenda helps in directing and managing the decision-makers' activities, which aids avoiding bias.

The methods addressed in chapter 2 all facilitate formal processes that provide structure and set agendas. Hence, they all satisfy the above stated requirement.

Dwarakanath (1996) mentioned the above requirement as 'decision support should help designers structuring the decision-making process'. Ullman and D'Ambrosio (1995) did not address it.

The higher time consumption of the two formal approaches was partly caused by the groups spending considerable time in developing general process approaches.

All three groups spent some time in discussing their general process approaches. This can be seen in tables 4.5, 4.6 and 4.7. The two formal groups spent 12% (formal students) and 9% (professionals) of their entire process durations on these discussions. In contrast, the informal students only spent 6% of their overall much shorter process on these discussions. Developing the formal approaches was by no means straightforward. Various suggestions were made and discussed in great length, but then found ineffective and abandoned. This was time consuming.

One of the formal groups, the formal students, spent much time at the beginning of their decision-making process in developing an approach and then worked through it without considerable deviations and with only little alterations. The other formal group, the professionals, spent less time at the beginning of their decision-making process in developing an approach, but discussed, altered and specified it frequently throughout the workshop. The result was a process that seemed to be altogether less efficient than the formal students' process. The engineers took substantially more time (80 minutes) than the students (60 minutes) to reach exactly the same conclusion, i.e. they selected the same alternative. This indicates that it may be of advantage to commence the process by specifying a clear approach that does not need to be altered later on during the process.

Requirement:

Decision-making support should offer pre-designed, formal structures for general decision-making processes that may be readily adopted by the decision-makers.

Benefit:

Reusing available process structures saves time otherwise required for developing them.

The methods addressed in chapter 2 all support formal structures that can be readily adopted for general decision-making processes. Hence, they all satisfy the above stated requirement.

Dwarakanath (1996) generally mentions the above requirement as 'decision support should present the designers with a quick overview of the decision-making process'. Ullman and D'Ambrosio (1995) did not address it.

Information capture

All groups generated valuable information during their decision-making processes. Yet, very little of this information was actually captured.

The informal students did not record any aspects of their decision-making process, whatsoever. This appears to be a reason for the group's low level of explicitness. That is, nothing forced them at any time to make a clear point that can be written down. I believe this causes difficulties in retrospectively comprehending the rationales underlying their decision - maybe even for themselves. I perceived these difficulties as soon as I tried to analyse the workshop, which was only a few days after it had been conducted.

The two formal groups recorded some particular aspects of their decision-making processes by taking notes. This made their processes explicit in regard to these aspects, which were criteria structures, criteria importance factors and formal evaluation scores. As stated in observation 6, the formal groups were absolutely consistent with respect to these aspects. This was in contrast to the informal group. Yet, observation 8 states that even the formal groups were actually only consistent with respect to those aspects that were recorded. Regarding other aspects, such as raising justifying evidence for evaluation scores, they were not consistent.

Another purpose of recording particular aspects of the decision-making process is helping the groups to remember what they have done already. This means, preventing them from forgetting their own discussions and reminding them about the exact conclusions of these discussions. Observation 12 highlights that repeated discussions of particular aspects occurred during all three groups' decision-making processes. There appeared to be at least three different reasons for repetitions. First, the group did not appear to be aware of having discussed exactly the same aspect before. Second, the group appeared to have forgotten the exact outcome of the earlier discussion. Third, the group needed to re-discuss an aspect in the light of newly gathered information.

Information reuse seems to be applicable in at least two situations. First, when the groups had come to the point when a decision could have been made, they looked for reassurance. The formal groups used their notes on the evaluation scores. They also needed some of the evidence justifying particular scores. But, even if evidence had been generated it was not recorded. Therefore it had to be recalled from memory, which will become ever more difficult the more time elapses between generation and recall.

Second, the groups indicated that further design activities, subsequent to the decision at hand, would concentrate on developing some identified issues of the selected alternative, taking into consideration the increased understanding of the existing alternatives, the product environment and the relevant evaluation criteria. That is, the further design process would actually be directed by these issues and related aspects. Therefore, capturing information on these aspects during the decision-making process helps in directing the further design process and prevents repeated information gathering of the same kind.

Requirement:

Decision-making support should be capable of capturing all relevant information that is generated during decision-making processes.

Benefits:

Capturing information aids the explicitness and consistency of the process, helps in monitoring the process and makes information available for reuse. This avoids repeated information gathering.

None of the methods addressed in chapter 2 satisfies the above stated requirement to the full extent. Even though most methods foster capturing some information, in particular criteria weights and evaluation scores, they do not aim at capturing all relevant information. This requirement seems to point in the direction of database/information-management systems.

Dwarakanath (1996) generally mentioned the above requirement as 'decision support should facilitate structured documentation of the decision-making process'. As an aim of his proposed decision support framework he suggested capturing all information used during the decision-making process. Ullman and D'Ambrosio (1995) generally addressed this requirement in their item 11, 'level of support: representation', which indicates a need for the representation of issues, alternatives, criteria and evaluations.

Within the following sections specific details will be discussed about exactly what sort of information should be captured as well as how it should be stored and processed.

5.3.2 Establishing a criteria structure

Within this section I will suggest and discuss particular aspects that, I think, should be addressed by decision-making support aiming at the effective handling of evaluation criteria. These aspects relate to (i) the meaning of criteria, (ii) the decomposition of criteria, (iii) criteria of general interest, (iv) types of criteria, (v) the flexibility of a criteria structure, and (vi) the relative importance of different criteria.

Meaning of criteria

All three groups based their evaluations, at least to some extent, upon a set of evaluation criteria. It appeared as if the meaning of these criteria was in many cases unclear or ambiguous. Different group members had a different understanding about what exactly particular criteria referred to.

When unclearness or ambiguities with respect to some criterion became evident, the groups started to discuss the meaning of this criterion as the activity 'defining criteria'. Such discussions were a core activity of the formal groups (see observation 2) and arose especially during times when criteria were structured and written down (see section 4.4.2, step 2B.1). This suggests that writing down the criteria in some form of a structure fosters discussions and reveals unclearness or ambiguity regarding the group's understanding of criteria.

The meaning of particular criteria were also repeatedly discussed throughout the decision-making process despite apparent earlier agreements. As no notes were ever taken regarding this aspect, I suggest that details about earlier agreements may have simply been forgotten by individual group members.

Requirement:
Decision-making support should be capable of storing the meaning of criteria, easily accessible at all times.
Benefits:
Storing the meaning of criteria aids clarity and avoids repeated discussions.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) generally mentioned the above requirement as 'decision support should facilitate capturing the criteria established'. Yet, I could not find any pointers towards the specific need for capturing and presenting the *meaning* of criteria in addition to listing them. This applies equally to Ullman and D'Ambrosio (1995).

Decomposition of criteria

The groups were given a set of evaluation criteria that was structured hierarchically. That is, criteria at higher levels were more abstract and general, whereas criteria at the bottom level were more specific and could be used directly to assess the alternatives' performances.

In particular the two formal groups spent considerable time in deliberating and changing this structure. At some 'branches' the given hierarchy was expanded and at others it was pruned (see section 4.4.2, step 2B.1.1). The apparent reason for expansion was the realisation that existing criteria were lacking particular metrics upon which to directly assess the alternatives. The apparent reason for pruning was the realisation that some metrics were too detailed for the evaluations at hand. It was already suggested in the previous section that the groups' efforts in structuring and re-structuring criteria triggered clarifying discussions about the criteria's meaning.

At the end of the workshops the groups discussed possible further stages in the design process. They suggested evolving the selected alternative thus creating a number of new, more detailed alternatives. These alternatives would then again be evaluated upon a similar set of criteria as the one applied for the original alternative's evaluation (see section 4.4.3, step 3.2). This implies that the developed criteria structure is reusable. Such reuse aids ensuring consistency throughout the evaluations of a particular design project.

Requirement:
Decision-making support should facilitate generating and storing hierarchically decomposed criteria structures.
Benefits:
Generating and storing criteria structures aids clarity and facilitates their later reuse.

Hierarchically decomposed criteria structures were addressed in chapter 2, in particular with respect to the Analytic Hierarchy Process (AHP) in sections 2.2.3 and 2.2.4. Žavbi and Duhovnik (1996) believe that the AHP's criteria structures enable a recording of knowledge on criteria, especially with respect to their relevance in the overall context. According to Salo and Hämäläinen (1997) this improves problem understanding and aids communication among decision-makers. My study supports their view. Some researchers even claimed that any decision-making problem with more than 7 ± 2 criteria needs to be arranged hierarchically in order to be at all manageable for human decision-makers (Jerčić & Bajić, 1993).

Also based on criteria hierarchies is Yang and Sen's (1997) method (see section 2.3.8). Hence, the AHP as well as Yang and Sen's (1997) method satisfy the above stated requirement. All other methods, as discussed in chapter 2, do not explicitly support this requirement.

Dwarakanath (1996) generally mentioned the above requirement by saying 'decision support should facilitate the structured documentation of criteria established'. More specifically, as part of his proposed decision support framework he suggested the generation of hierarchical criteria structures. Ullman and D'Ambrosio (1995) did not address this requirement.

Criteria of general interest

The criteria given to the groups had a specific relevance to the particular design issue at hand. This also means that these criteria were rather specific for those alternatives being evaluated.

The informal students as well as the professionals identified additional criteria. Yet, there were two interesting differences between the informal students and the professionals regarding this aspect: First, the professionals intentionally thought about additional criteria (see section 4.4.2, step 2B.1.1), whereas the students found them unintentionally, as a by-product of their decision-making process (see section 4.4.2, step 2A). Second, the professionals identified entirely new criteria, whereas the students identified only new metrics, related to given criteria. The professionals' new criteria were of general relevance, beyond the particular design issue at hand and the specific alternatives being evaluated, such as 'environmental friendliness' and 'ease of product development'. No such general thoughts were raised by any of the students.

Requirement:
Decision-making support should raise awareness for criteria of general interest.
Benefit:
Being made aware of such criteria may prevent in particular inexperienced decision-makers from ignoring them.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) did not directly mention the above requirement. Yet, as part of his proposed decision support framework he suggested that some criteria may be predefined. Such predefined criteria may be of general interest. Ullman and D'Ambrosio (1995) did not address this requirement.

Types of criteria

The formal students explicitly distinguished their evaluation criteria into objectives and constraints. They differentiated these two types of criteria according to their ability to compensate for each other in trade-offs. Objectives do compensate and constraints do not. This means, if an alternative fails to satisfy a single constraint, the entire alternative was rendered unacceptable, regardless of its performance with respect to other criteria. Also, in contrast to an objective, a constraint was seen as being binary, i.e. it could either be satisfied or not.

Although the other groups did not explicitly distinguish between objectives and constraints, there were indications that, at least the professionals perceived such distinction as well (see section 4.4.2, step 2B1.1). The professionals discussed whether they should introduce a 'constraint boundary' on a criterion that was otherwise seen as an objective (the professionals treated all criteria in a compensating manner, like objectives rather than constraints). They indicated, if this boundary (particular performance for criterion 'time of each delivery – see table 4.2) was exceeded, the entire alternative would be useless, regardless of its other performances. If not exceeded, the objective was to have a time as low as possible, which could be traded-off in a compensating manner among other objectives. This means that the professionals actually constrained an objective.

Requirement: Decision-making support should facilitate modelling the criteria as objectives, constraints and constrained objectives. Benefit: The ability to model these types of criteria allows for a realistic representation of decision-making situations.

Constraints and objectives were addressed in chapter 2, with particular respect to Multi-Attribute Value Theory (MAVT) in sections 2.2.1 and 2.2.2. Neither MAVT nor any of the other methods as discussed in chapter 2 actively support differentiating criteria into constraints, but none of them seems to prohibit such differentiation.

The notion of a constrained objective was discussed in chapter 2, with respect to the outranking methods, in sections 2.2.5 and 2.2.6. Especially when using Electre, a so called 'veto threshold' may be defined on any criterion. This threshold indicates the point at which an alternative is so outperformed that it needs to be rejected regardless of its performance with respect to other criteria (Simpson, 1996). This is what the professionals suggested when they constrained the (otherwise objective) criterion 'time for each delivery'. Hence, the outranking methods applying a veto threshold are particularly supportive for modelling constrained objectives and therefore best satisfy the above stated requirement.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

Flexibility of the criteria structure

The criteria structure as given to the groups and as further developed by them were paper-based. That means, they were drawn on a sheet of paper.

I observed that the criteria space remained dynamic throughout the decision-making process: the informal students as well as the professionals identified additional criteria, not only before the evaluations took place, but also as part of the evaluations (see section 4.4.2, figures 4.13 and 4.18). The informal students never updated the given criteria structure. The professionals altered it and re-drew it before the evaluations took place. Yet, after it had been re-drawn it was not updated any more during the evaluations. This was despite the identification of a very relevant new criterion ('ease of product development').

Altering the paper-based structures was a lot of effort, as they had to be re-drawn. For the professionals this would have also meant re-drawing their entire evaluation matrix. I suggest the reason for the groups failing to update their structures was that they just did not want to make the effort. This would imply that paper-based structures were too rigid for coping with the obviously dynamic criteria space. The result was loss of information because newly identified criteria were not written down or captured in any way for reuse later on in the design process.

Requirement:

Decision-making support should offer more flexibility for storing and presenting criteria structures than paper-based decompositions.

Benefits:

Flexibility with respect to the criteria structure allows for easy alterations, in particular the inclusion and exclusion of criteria at any time. This aids information storage and consistency.

The methods addressed in chapter 2 may all satisfy the above stated requirement if applied e.g. as flexible computerised application.

Dwarakanath (1996) did not directly mention the above requirement. Yet, his proposed decision support framework suggests flexible computational support for capturing, storing and presenting criteria. Ullman and D'Ambrosion (1995) addressed it in their item 1, 'problem completeness: incomplete', indicating that decision support needs to cope with evolving sets of criteria.

Relative importance of different criteria

Observation 5 stated that only the formal students explicitly communicated the relative importance of criteria. However, I argued that there were indications for, in particular, the informal students also perceiving some of the criteria as more important than others, even though this remained implicit.

Requirement:

Decision-making support should encourage communicating the perceived importance of the different evaluation criteria.

Benefit:

Communicating the importance of criteria aids the decision-making process' explicitness.

Most of the methods for decision-making support, as discussed in chapter 2, involve communicating the importance levels of evaluation criteria. An exception is Pugh's (1990) method of controlled convergence. As discussed in sections 2.2.5 and 2.2.6, when using this method, criteria importance does not need to be considered. Hence, the method of controlled convergence does not satisfy the above stated requirement, whereas the other methods do so.

Dwarakanath (1996) did not directly mention the above requirement. Yet, as part of his proposed decision support framework he acknowledged criteria importance being a required input for decision-making methods. Ullman and D'Ambrosio (1995) addressed the requirement in their item 4, 'objective function', which involve criteria importance values, and their item 5, 'consistency', which indicates that different decision-makers may communicate different perceptions about criteria's importance.

Aspects that need to be considered in association with making the perceived importance of criteria explicit are (i) the assignment of weightings, (ii) the format of weightings, and (iii) the justification of weightings.

Assignment of weightings

The two formal groups used numeric functions to aggregate information on an alternative's performance. This was done by combining information on individual criteria (restricted performance) into information on an alternative's overall (comprehensive) performance (see section 4.4.3, step 3.1). For such aggregations, criteria weightings may be used as factors that reflect the decision-makers' values, i.e. the importance of criteria.

The formal students did use such factors, whereas the professionals did not use such factors. They quite clearly communicated that the assignment of criteria weightings was left out for simplicity reasons. It was considered too time consuming (see section 4.4.2, steps 2B.1.2 and 2B.1.3).

The problem of using a formal aggregation function without criteria importance factors is that the results might be rendered meaningless. This is because if no criteria importance factors are assigned, the formal aggregation function automatically assumes that all criteria are of equal importance. Thus, only marginally important criteria have a far too high influence on the calculated comprehensive performances, whereas highly important criteria have far too little influence on the calculated comprehensive performances.

Requirement:
Decision-making support should make it easy and straightforward for the decision-makers to express their values as criteria importance factors (weightings).
Benefit:
Making it easy for the decision-makers to assign criteria importance factors encourages them to do so and thus aids in achieving meaningful results.

The assignment of criteria importance factors (weightings) was addressed in chapter 2 with respect to most methods of multi-criteria decision-making as discussed in section 2.2. Generally speaking, Ahn and Dyckhoff (1997) urged that for practicality reasons it must be ensured that the decision-making process is not made unnecessarily difficult when using a method. This view supports the above stated requirement. Yet, it does not seem to be absolutely clear how a method should elicit meaningful criteria importance factors without making the process unnecessarily difficult.

Many researchers agree that in practice, assigning meaningful criteria importance factors is rather difficult (Tebay et al., 1984; Brans et al., 1986; Roy, 1991; Simpson, 1996). Otto and Antonsson (1991) argue that it is not usually possible for decision-makers to a priori specify these factors; instead they set preliminary estimates and gain insight through iteration. I have not observed this in my study, but I had the impression that my groups were not very serious about specifying meaningful weightings (see observation 5).

Criteria importance factors may be assigned directly or in a comparative manner. Direct assignments are implied by most decision-making methods, such as Multi-Attribute Value Theory (MAVT) and the different outranking methods. However, it has been argued that direct weight assignments are too abstract for decision-makers and may result in inaccuracies (Zahedi, 1986). In contrast, comparative weight assignments are claimed to be natural and therefore appealing for practical decision-making (Sen & Yang, 1995; Weiss & Hari, 1997; Pöyhönen & Hämäläinen, 2000). In my study a mixed approach was taken: sometimes comparative and at other times direct (see figure 4.8).

A method that supports criteria weight assignments through iteration and by comparisons rather than by direct assignment is the Analytic Hierarchy Process (AHP) as discussed in sections 2.2.3 and 2.2.4. Hence, the AHP appears to satisfy the above stated requirement. Yet, although the AHP's *general* approach towards weight assignments, i.e. iteration and comparisons, appears easy and straightforward, the *specific* procedures necessary are sometimes considered as being difficult or tedious (Ghotb & Warren, 1995; Sen & Yang, 1995; Naudé et al., 1997). This means, actually not even the AHP seems to fully satisfy the above requirement.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

Formats of weightings

The group who assigned weightings, i.e. the formal students, applied a numeric aggregation function that required the criteria importance factors to assume a particular format. The factors had to be precise, cardinal numbers. Any other format could not have been processed by the applied function. However, this format did not seem to be very natural (see section 4.4.2, steps 2B.1.2 and 2B.1.3).

The group often assessed criteria-importance on a qualitative basis (see table 4.9) without indicating any *extent* of importance difference between criteria. This can only lead to ordinal weightings, which cannot be used as factors in numeric aggregation functions. The group was also capable of expressing the extent of importance differences between criteria, but they never did so in a precise, quantitative manner. The closest natural format was linguistically vague quantifications. Yet, the vague quantifications were then transformed into precise quantifications (see section 4.4.2, step 2B.1.3) to gain the required factors. This implies that the factors are somewhat questionable because a level of precision is pretended that was not inherent in the original, linguistically vague information.

Requirement:

Decision-making support should be capable of processing vague, linguistic expressions as a valid format for criteria-importance factors (weightings).

Benefit:

The capability of processing vague, linguistic expressions enables the decision-makers to express themselves naturally without the need for distorting transformations.

Formats of criteria importance factors were addressed in chapter 2, with respect to almost all methods for multi-criteria decision-making as discussed in section 2.2 and also with respect to methods for modelling uncertainty as discussed in section 2.3.

Most of the methods addressed in chapter 2 process precisely quantitative weighting factors. This is regardless of whether the factors are elicited directly, as in Multi-Attribute Value Theory (MAVT) or through comparisons, as in the Analytic Hierarchy Process (AHP). The only exception are methods that apply fuzzy set theory, as discussed in sections 2.3.11 and 2.3.12. Hence, none of these methods apart from the fuzzy ones seem to satisfy the above stated requirement.

Neither Dwarakanath (*1996*) nor Ullman and D'Ambrosio (*1995*) mentioned or addressed the above stated requirement.

Justification of weightings

Another aspect that may have rendered some of the group's criteria importance factors questionable is an apparent lack of justification. Even though for most, but not all, criteria weightings some justifying evidence was raised (see section 4.4.2, step 2B1.3), observation 5 suggests that it seemed to be a formality.

The group had prepared a form (their evaluation matrix) that included a specific space to be filled with the criteria importance factors. However, there was no space in this form to be filled with justifications for these factors. I suggest that the group's inconsistency and lack of thoroughness regarding the justifications was caused by that missing space. This is because there was nothing that forced the group to justify their factors in a consistent and thorough manner.

Requirement:

Decision-making support should include a feature that prompts the decision-makers to associate their criteria-importance factors with an effective justification.

Benefits:

Prompting encourages decision-makers to gather and communicate evidence. This, in turn, aids avoiding little consistency and little thoroughness with respect to criteria importance factors.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

5.3.3 Handling the alternatives

Within this section I will suggest and discuss particular aspects that, I think, should be addressed by decision-making support aiming at the effective handling of alternative design solutions. These aspects relate to (i) the presentation of alternatives and (ii) the development of alternatives.

Presentation of alternatives

All groups repeatedly clarified the working principles of the alternative design solutions (see section 4.4.2, step 2A and step 2B.2A/B.3). This may suggest an insufficient quality of the sketches and explanations provided by the researcher.

Understanding the alternative design solutions can be seen as a prerequisite for any meaningful evaluation. I observed partly repeated discussions that aimed at gaining such an understanding, particularly when some alternative was being evaluated. That is, when it became obvious that there was a lack of understanding among the group. Comments by some group members usually generated clarity and a unified understanding across all group members (see section 4.3.1, activity 'clarifying concept working principles'). None of these clarifying comments were ever written down.

Requirement:

Decision-making support should facilitate the capture of clarifying comments on the alternatives' working principles.

Benefits:

Capturing additional comments on the alternatives' working principles aids clarity and helps to avoid repeated clarifying discussions.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) generally mentioned the above requirement by saying that decision support should (i) present designers with all the alternatives proposed, (ii) facilitate capturing and organising documentation on all the alternatives proposed and (iii) provide an easy and flexible way of retrieving captured information. As part of his proposed decision support framework he suggests that information on alternatives should be captured in textual and graphical form. Ullman and D'Ambrosio (1995) generally addressed this requirement in their item 11, 'level of support: representation', which indicates that decision-making support involves representing information on alternatives.

Development of alternatives

All groups further developed the given alternatives. Such developments referred to sub-issues (see section 4.3.1, activity 'deliberating sub-issues'). There are two aspects to be considered: first, relations between issue and sub-issue and second, sub-issue solutions.

Relations between issue and sub-issues

The groups' deliberations on sub-issues involved (i) identifying the issue/sub-issue relationship, i.e. the sub-issue's implication for the evaluation at hand; (ii) generating solutions for the sub-issue; and (iii) evaluating the solutions.

After the informal students had established their overall ranking order, extensive deliberations focused on possible solutions for two identified sub-issues, which were associated with the highest ranked alternative only (see section 4.4.3, step 3.2). One of the sub-issues had been deliberated in great length before. There was no apparent reason for the repetition at this point apart from the group having forgotten about the details of their earlier deliberations.

The two formal groups in particular discussed the further design stages when they had completed their decision-making process. The student group addressed a specific sub-issue, which seemed to be a weak aspect of the highest ranked alternative. Addressing this aspect appeared to be an immediate design task that aimed at improving and further detailing the alternative. This implies that the further design activities are directed towards the identified sub-issues. The professionals explicitly suggested that they would evolve the identified sub-issues of the highest ranked alternative into more detail, thereby creating a number of alternatives based on the one ranked highest. Therefore, it seems worthwhile to make notes on what sub-issues have been identified during the alternatives' evaluations. Yet, none of the groups did so.

Requirement:
Decision-making support should model any identified issue/sub-issue relationships.
Benefits:
Modelling issue/sub-issue relationships helps to prevent decision-makers from 'forgetting' them during the decision-making process, and helps to direct the further design process.

None of the methods addressed in chapter 2 satisfies this requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) mentioned the above requirement by saying that 'decision support should facilitate presenting sub-issues'. As part of his proposed decision support framework he suggested arranging issue/sub-issue relationships as issue hierarchies. Ullman and D'Ambrosio (1995) addressed it in their item 10, 'range of issue independence', which indicates the existence of different relationships between issues, including sub-issues.

Sub-issue solutions

The outcome of deliberating a sub-issue is that the group accepts a solution for it. Such a sub-issue solution is then assumed to be part of the main alternative (see section 4.3.1, sub-activity 'accepting assumptions about sub-issue solutions'), which can thus be seen as further developed. Therefore, this sub-issue solution becomes a basis for the main alternative's evaluation, which in turn can then only be appreciated retrospectively if the underlying sub-issue solution is known.

I observed that in many cases the deliberations on sub-issues were terminated without explicitly accepting any solution (see section 4.4.2, step 2B.2A/B.3, figure 4.19). However, the implication of most sub-issues for the decision-making process was that some solution had to be found to proceed with the evaluation. What the groups often seemed to do was implicitly accept a sub-issue solution: individual group members appeared to believe that it was understood by the entire group that some mentioned solution was selected and accepted to be part of the main alternative. This introduced ambiguity to the process because it was not quite clear whether other group members actually had the same understanding.

Requirement:

Decision-making support should capture an associated solution for identified issue/sub-issue relationships.

Benefit:

Capturing the associated solutions for issue/sub-issue relationships helps to avoid ambiguity.

None of the methods addressed in chapter 2 satisfies this requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) mentioned the above requirement by saying that 'decision support should facilitate identifying unresolved issues'. As part of his proposed decision support framework he suggested that issues which have not been resolved should be highlighted. Ullman and D'Ambrosio (1995) generally addressed this requirement in their item 10, 'range of issue independence', which indicates the need for resolving sub-issues before 'super-issues' may be resolved.

5.3.4 Describing the product environment

Within this section I will suggest and discuss particular aspects that, I think, should be addressed by decision-making support aiming at the effective handling of descriptions on the product environment. These aspects relate to (i) the known environment and (ii) the assumed environment.

Known environment

The evaluation of an alternative design solution is to a large extent related to how a product, based on the design, would perform in a given environment. Therefore, a clear understanding of this environment is vital for any meaningful evaluation. All groups tried to determine aspects of the product environment at some point during their decision-making processes (see section 4.3.1, activity 'clarifying the product environment' and section 4.3.3, figure 4.4). This may suggest that the handouts given to the groups were inadequate in describing the product environment.

In some cases the groups discussed the product environment on the basis of what was stated in the handouts. That is, the discussions sought clarification and a unified understanding of the environment across all group members. Such discussions resulted in clarifying amendments to the handouts, which were, however, not written down.

In other cases the groups revealed that the given handouts were not complete. The groups would then try to gain external information. I tried to provide this information by using additional material or knowledge that had previously not been considered useful for the workshops. Yet again, written notes were not made.

Requirement:
Decision-making support should store any information on the product environment generated during the decision-making process.
Benefits:
Capturing information on the product environment aids clarity and helps to avoid repeated information gathering.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) did not directly mention the above requirement. Yet, as part of his proposed decision support framework he suggested capturing all information used during the decision-making process. This should include information on the product environment. Ullman and D'Ambrosio (1995) did not address this requirement.

Assumed environment

Occasionally particular aspects of the product environment could not be determined. This was due to lack of information. To proceed with the evaluations the groups would then work with assumptions (see section 4.3.1, sub-activity 'making assumptions on the product environment', section 4.4.2, step 2B.2A/B.3, figure 4.18).

Assumptions need to be stored because evaluations are made on the basis of them. This means that if a particular evaluation needs to be understood retrospectively it is necessary to know of any underlying assumptions. However, assumptions should not be confused with what is known. In immediate further activities assumptions need to be verified. This, in turn, verifies the evaluations made upon them.

Requirement:

Decision-making support should store assumptions on the product environment, but keep them distinct from what is known about the environment.

Benefits:

Storing assumptions distinct from what is known avoids confusion and directs further activities.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

5.3.5 Evaluating the alternative design solutions

Within this section I will suggest and discuss particular aspects that, I think, should be addressed by decision-making support aiming at an effective evaluation of the alternative design solutions. These aspects relate to (i) the evaluation approach, (ii) the evaluation basis, (iii) the applied evaluation scales, (iv) justifying evidence, (v) confidence and (vi) the establishment of comprehensive rankings.

Evaluation approach

All groups first spent some time on the evaluation of the individual alternatives in process step 2 before an overall (comprehensive) ranking order across all alternatives was established in process step 3. There were two general approaches towards the alternatives' evaluation: an informal 'comprehensive approach' and a formal 'criteria-based approach' (see section 4.4.2).

Observation 6 stated that the informal comprehensive approach was characterised by little explicitness, high level of inconsistency and bias. This was in contrast to the formal criteria-based approach.

Requirement:
Decision-making support should be built upon a criteria-based evaluation approach.
Benefits:
The criteria-based approach aids explicitness as well as consistency and helps to avoid bias.

Principal evaluation approaches were addressed in chapter 2, with respect to the methods for multi-criteria decision-making as discussed in section 2.2. Multi-Attribute Value Theory (MAVT) and the Analytic Hierarchy Process (AHP) decompose complex multi-criteria decision problems into a set of single criterion sub-problems, which are systematically resolved and then aggregated to obtain a solution to the larger decision problem (Weiss & Hari, 1997). Although not in an axiomatic manner as in MAVT and AHP, but more in qualitative algorithmic form, this approach is also taken by the outranking methods (Simpson, 1996). This implies that all methods discussed in section 2.2 use the criteria-based approach. The same applies for the methods discussed in section 2.3. That is, Multi-Attribute Utility Theory (MAUT), the

Engineering Decision Support System (EDSS), Yang and Sen's (1997) method as well as the fuzzy methods are all explicitly based on evaluation criteria. Hence, they all satisfy the above stated requirement.

Dwarakanath (1996) mentioned the above requirement by saying: decision support should facilitate checking whether all alternatives' evaluations have been based upon all criteria. Ullman and D'Ambrosio (1995) addressed this requirement in their item 4, 'objective function', which indicates the use of criteria-based aggregation functions, and in their item 6, 'comparison basis', which considers how criteria may be used for evaluating alternatives.

Those groups who used the formal criteria-based approach took considerably more time to reach their decisions than the group who used the comprehensive approach (see section 4.2). One of the reasons was that the formal groups spent a lot of time in designing their approaches through the activity 'discussing the process approach' (see section 4.3.3, figure 4.4). Another reason was that, due to their consistency, they actually evaluated more alternative-criterion pairs than the informal students.

The professionals did not distinguish the criteria according to different levels of importance. As a result, they evaluated all alternatives with respect to all criteria, which was time consuming. The formal students did distinguish the criteria according to different levels of importance. But, they too evaluated all alternatives with respect to all criteria and seemed to dedicate the same effort for each criterion, even for the least important one.

Requirement:
Decision-making support should facilitate focusing the evaluations towards the important criteria.
Benefit:
Focusing on the important criteria avoids spending time on criteria that should have little influence on the final decision, without being inconsistent or biased.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

Evaluation matrix

Both formal groups applied evaluation matrices to implement their criteria-based approaches. These matrices generated alternative-criterion pairs by associating each alternative with each criterion (see section 4.2, tables 4.1, 4.2, 4.3).

The matrices seemed to be the main factor in making the formal groups' decision-making processes more explicit than the informal students' process. This is because the matrices forced the groups to express their evaluations in a specific, formal manner and write them down.

The matrices ensured clarity and consistency regarding the application of evaluation criteria. By clearly associating each alternative with each bottom level criterion (metrics) it was ensured that the same set of criteria was used for the evaluation of each alternative and that always the bottom level criteria, i.e. the metrics were applied.

When the formal groups analysed their established rankings in process step 3.2, there were frequent consultations on the completed evaluation matrices. As the matrices contained the evaluation scores for all alternative-criterion pairs, they reflected, at least to some extent, the groups' rationale.

Requirement:

Decision-making support should apply evaluation matrices.

Benefits:

Applying evaluation matrices aids explicitness as well as clarity, helps ensuring consistency in complex situations and facilitates the retrospective decision analysis.

It was said above that all of the methods addressed in chapter 2 apply the criteria-based approach. I suggest that this means that they may all be laid out in such a way that evaluation matrices could be established. Hence, they are all capable of satisfying the above stated requirement.

Dwarakanath (1996) generally mentioned the above requirement by saying that 'decision support should automatically create inputs for evaluation methods'. More specifically, as part of his proposed decision support framework he suggests using evaluation matrices. Ullman and D'Ambrosio (1995) did not address this requirement.

Flexibility of the matrix

The applied evaluation matrices were paper-based. That is, they were drawn by hand on a sheet of paper. This means that it was difficult to change what had been written down.

Observation 12 stated that during all groups' decision-making processes there were iterations. These seemed to be caused by a number of reasons. First, when the groups evaluated some alternative-criterion pair they occasionally realised that a previous evaluation of another alternative with respect to the same criterion had to be changed relative to the current evaluation. Second, when the groups gained new insight in some alternative's working principle it was realised that previous evaluations were based on a wrong understanding of this alternative's working principle and had therefore to be repeated. Third, when the groups gained new insight in the product environment it was realised that particular previous evaluations were based on a wrong understanding of the environment and therefore had to be repeated. In short, the generation of new information on other alternatives' performances, on alternatives' working principles and on the product environment occasionally made it necessary to repeat earlier evaluations. However, the paper-based evaluation matrices made it difficult to change input without jeopardising clarity.

Requirement:
Decision-making support should offer more flexibility for evaluation matrices than can be achieved by simply using paper.
Benefit:
Flexible matrices allow for changing input at any time.

The methods addressed in chapter 2 may all satisfy the above mentioned requirement if implemented e.g. as a flexible software application.

Dwarakanath (1996) did not directly mention the above stated requirement. Yet, his proposed decision support framework suggests flexible computational support for capturing, storing and presenting *evaluation information*. Ullman and D'Ambrosion (1995) addressed it in their item 1, 'problem completeness: incomplete', indicating that decision support needs to cope with evolving information on the 'decision space', i.e. criteria and alternatives.

Evaluation basis

In observation 7 it was stated that, in particular, the formal groups had a natural tendency to compare alternatives rather than to evaluate them on an independent, absolute basis. It seemed as if it was impossible for the groups to evaluate on an independent, absolute basis, even when they tried.

Comparative evaluations rank alternatives, but do not clearly state whether any of them are sufficiently effective on an independent, absolute scale. Therefore, the question on whether or not even the highest ranked alternative is worth further detailing or should be abandoned cannot be clearly answered based on such evaluations. This was reflected by the groups' answers when I asked this question (see section 4.4.3, step 3.2). However, as the groups did provide answers, even though not clearly based on their evaluation scores, they seemed to have gathered an 'impression' through the foregone exploration and evaluation of the various alternatives. It appears therefore not necessary to force the decision-makers into attempting independent, absolute evaluations. Instead, 'information rich' comparative evaluations seem to foster sufficient 'learning' about the decision-making situation at hand.

In addition to the comparative evaluations, it would be beneficial for directing further design activities if the decision-makers qualitatively marked problem *areas*, i.e. criteria for which even the highest ranked alternative seems to perform less than desirable (if it is know what 'desirable' in fact means).

Requirement:

Decision-making support should be based on comparative evaluations. But decision-makers should be prompted to mark problem areas (criteria).

Benefits:

Comparative evaluations are natural. Prompting for problem areas aids directing the further design activities.

Comparative evaluations were addressed in chapter 2, with respect to the Analytic Hierarchy Process (AHP) and the outranking methods. It was claimed that the AHP's comparisons, as opposed to absolute evaluations, describe natural decision-making behaviour in many situations and are therefore considered appealing for practical decision-making (Sen & Yang, 1995; Weiss & Hari, 1997; Pöyhönen & Hämäläinen, 2000). This was particularly emphasised for decision-making in conceptual engineering design, as there are usually no standard scales or absolute measures available with which the expected performances of a concept may be assessed on an independent, absolute basis (Chen & Lee, 1993; Žavbi & Duhovnik, 1996; Weiss & Hari, 1997). This has become evident in my study too (see observation 7).

In common with the AHP, the outranking methods are based on the assumption that a decision-maker can only express their evaluations by comparisons (De Keyser & Peters, 1996). All other methods, as discussed in chapter 2, are based on independent, absolute comparisons. Hence, only the AHP as well as the outranking methods satisfy the above stated requirement.

Dwarakanath (1996) did not mention the above requirement. Ullman and D'Ambrosio (1995) addressed it in their item 6, 'comparison basis', which discusses the existence of absolute as well as comparative evaluations.

Variants

There were two variants for the evaluation of alternatives in process step 2B.2: variant A, 'criterion-alternatives' and variant B, 'alternative-criteria' (see section 4.4.2, figure 4.17). As stated in observation 7, if a number of alternatives are available, variant A seems very natural and almost inevitable for comparative evaluations. This means, first picking a criterion and then evaluating all alternatives upon this criterion before picking another criterion.

> Requirement:
> Decision-making support should be based on variant A, 'criterion-alternatives' for restricted evaluations.
> Benefit:
> Variant A, 'criterion-alternatives' is naturally applied by decision-makers for comparative evaluations and therefore is straightforward and convenient.

All of the methods addressed in chapter 2 can be applied in such a way that evaluations according to variant A, 'criterion-alternatives' are implemented. Hence, they are all capable of satisfying the above stated requirement.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

Evaluation scales

Observation 7 stated that the groups used two general types of scales for their evaluations: binary or non-binary. The informal students' advantage/disadvantage gathering as well as the formal students' checking for constraint satisfaction represented binary evaluations. All other evaluations were non-binary, i.e. the groups were able to express various degrees of satisfaction. Essentially I see such degrees of satisfaction as valuable information that helps distinguishing the alternatives' performances from each other. Obviously, this distinctive information is lost if the applied evaluation scales are binary.

> Requirement:
> Decision-making support should model various degrees of satisfaction regarding performance evaluations.
> Benefit:
> The possibility to express degrees of satisfaction helps to prevent loss of information.

Strictly binary evaluations were addressed in chapter 2, with respect to Herling's (1997) Engineering Decision Support System (EDSS) as discussed in section 2.3.6. The EDSS rests upon a boolean (binary) position. This means that the evaluations are about whether an alternative will satisfy a criterion rather than about the degree to which satisfaction is achieved (Ullman et al., 1997). As such, EDSS does not satisfy the above stated requirement.

Some of the outranking methods, as discussed in sections 2.2.5 and 2.2.6 compare alternatives on the basis of the 'quasi-criterion' (Brans et al., 1986; Roy, 1991), also called '0-1 criterion with indifference area' (Huylenbroeck, 1995). These are, in particular, decision-making with QFD, as described by Hales (1995), and Pugh's method of controlled convergence. Methods using the 'quasi criterion' are not strictly binary. They allow for evaluations resulting in 'better than', 'as good as' or 'worse than' statements for pairwise comparisons. Even though not strictly binary, these evaluations cannot express more than two degrees of satisfaction regarding performance evaluations either. Hence, I consider these methods as not satisfactory regarding the above stated requirement. All other methods addressed in chapter 2 facilitate expressing degrees of satisfaction and therefore satisfy the requirement.

Dwarakanath (1996) did not mention the above requirement. Ullman and D'Ambrosio (1995) addressed it in their item 7, 'dimension', saying that decision-makers express *how well* an alternative meets a criterion, which suggests non-binary evaluations.

Scoring formats

The two formal groups applied evaluation scales that either contained qualitative scores or precisely quantitative scores. Whereas the qualitative scores seemed to largely match the groups' natural behaviour, the precisely quantitative scores did not (see section 4.4.2, step 2B.2A/B.3). That is, the groups naturally (informally) either expressed themselves qualitatively or vague quantitatively, but rarely in a precise quantitative manner (see section 4.3.4, figure 4.9).

The groups often expressed evaluations on a qualitative basis without indicating any *extent* of satisfaction difference between the alternatives' performances. This can only lead to ordinal scores, which cannot be used as factors in numeric aggregation functions as applied by the groups. Yet, the groups were also capable of expressing the extent of satisfaction differences, but these were usually expressed in imprecise, linguistically vague terms. By mapping such terms onto qualitative or precise quantitative scales the original meaning of any linguistically vague quantifications was distorted (see section 4.4.2, step 2B.2A/B.3). This means that information was either lost or artificially enhanced by pretending a level of precision that was not inherent in the original vague quantification.

Requirement:

Decision-making support should be capable of processing linguistically vague quantifications for expressing any extent of satisfaction differences between the alternatives' performances.

Benefit:

The capability of processing linguistically vague quantifications aids the adequate representation of available information without transformations.

With one exception, none of the methods as discussed in chapter 2 are capable of processing linguistically vague quantifications for expressing extents of satisfaction differences. Most methods are either based on precise quantitative assessments, such as Multi-Attribute Value Theory (MAVT) and the Analytic Hierarchy Process (AHP), or on pure qualitative assessment, as the outranking methods. The exception are methods that apply fuzzy set theory, which, as discussed in sections 2.3.11 and 2.3.12, can process vague quantifications. Hence, none of the methods apart from the fuzzy ones satisfy the above stated requirement.

If, through gathering additional information and advances in the design process, more precise quantitative assessments become available, these can still be modelled and processed using fuzzy mathematics, i.e. there is mathematical continuity between vague quantifications and precise quantifications (Allen, 1996). This means even though precision will increase through information gathering, the mathematical formulation of this information, regardless of whether it is precise or vague, remains the same if it is based on fuzzy sets. Hence, fuzzy methods also satisfy the above stated requirement with emphasis on expressing *any extent* of satisfaction differences, even if it is a precise quantification.

Dwarakanath (1996) did not mention the above requirement. Ullman and D'Ambrosion (1995) addressed it in their item 2, 'abstraction', which indicates the occurrence of qualitative, quantitative and mixed evaluation information, and in their item 3, 'determinism', which discusses point values and inexact distributions regarding evaluation information.

Evidence

Observation 8 stated that even though the formal groups were consistent regarding the evaluation of all alternatives with respect to all criteria they were not consistent with respect to explicitly providing evidence to justify these evaluations (see also section 4.4.2, step 2B.1.3). This rendered a number of evaluations being not explicitly justified and therefore appearing somewhat unclear, which makes it difficult to retrospectively comprehend them.

The groups had prepared a form (their evaluation matrix) that included a specific space to be filled with the evaluation scores. However, there was no space in this form to be filled with justifications for these scores. I suggest that the groups' inconsistency regarding the justifications was caused by that missing space. Hence, there was nothing that forced the groups to provide justifying evidence in a consistent manner.

Requirement:

Decision-making support should include a feature that prompts the decision-makers to associate their evaluation scores with justifying evidence.

Benefits:

Prompting encourages the decision-makers to gather and communicate evidence. This, in turn, increases the decision-making process' explicitness, clarity and ease of retrospective analysis.

None of the methods addressed in chapter 2 satisfies the above stated requirement, which appears to refer to a specific aspect of an information-management system.

Dwarakanath (1996) did not directly mention the above requirement. Yet, as part of his proposed decision support framework he suggested that the reasons behind why particular alternatives were rejected should be captured. Such reasons may represent justifying evidence. Ullman and D'Ambrosio (1995) did not address this requirement.

Confidence

Observation 10 stated that the groups occasionally expressed the likelihood of some evaluation being correct. The groups never expressed their confidence regarding evaluations as numeric probability in percent, but always by using linguistic terms. I see these expressions as valuable information, enriching the evaluation scores.

Requirement: Decision-making support should be capable of modelling the decision-makers' confidence in evaluations. It should allow for linguistic terms rather than numeric expressions. Benefit: Capturing confidence, especially in linguistic terms, helps towards modelling complete information on the decision situation.

Chapter 2 addressed three methods for expressing likelihoods regarding the alternatives' performances: probability theory (section 2.3.1), Bayesian inferencing (section 2.3.4) and Dempster-Shafer theory (section 2.3.7).

Probability theory models random events (Mizrahi & Sullivan, 1993). Randomness is linked with the occurrence of chance phenomena. Yet, the confidence in evaluations expressed by the groups did not describe chance phenomena, but subjective belief which, in turn, cannot be modelled by probability theory.

Subjective belief, rather than objective probabilities, can be modelled by Bayesian inferencing and the Dempster-Shafer theory (Graham & Jones, 1988). The difference between these two methods is that the former cannot model 'ignorance', whereas the latter can (see section 2.3.7). Ignorance, expressing that the groups do not know any likelihood, was shown whenever evaluations were made without any confidence assignments. This means, if no confidence was expressed for some evaluation, the group did not try to say that they have little confidence, but instead they did not know, i.e. they were ignorant. Hence, the Dempster-Shafer theory appears most appropriate for modelling confidence according to the above stated requirement. Yang and Sen (1997) demonstrated that this method can be applied in connection with linguistic terms for expressing confidence.

Dwarakanath (1996) mentioned the above requirement by saying that decision support should facilitate showing weak and strong points of an alternative by presenting the confidence of designers in their evaluations. Ullman and D'Ambrosion (1995) addressed the requirement in their item 7, 'dimension', which indicates the occurrence of expressions on the decision-makers' confidence regarding evaluations.

Comprehensive rankings

Both formal groups comprehensively ranked the alternatives by applying numeric aggregation functions. Applying such functions involved calculations, which were represented by the activity 'evaluating comprehensive performances formally'. This activity took effort and considerable time (see section 4.3.3, sub-activity figures 4.6 and 4.7).

Requirement:
Decision-making support should offer a feature for automating any calculations necessary for applying aggregation functions.
Benefit:
Automating calculations for aggregation functions facilitates time savings.

The methods addressed in chapter 2 may all satisfy the above stated requirement if implemented e.g. as a flexible software application.

Neither Dwarakanath (1996) nor Ullman and D'Ambrosio (1995) mentioned or addressed the above requirement.

A summary of all requirements for decision-making support as identified in my study and as discussed above is given in figure 5.1.

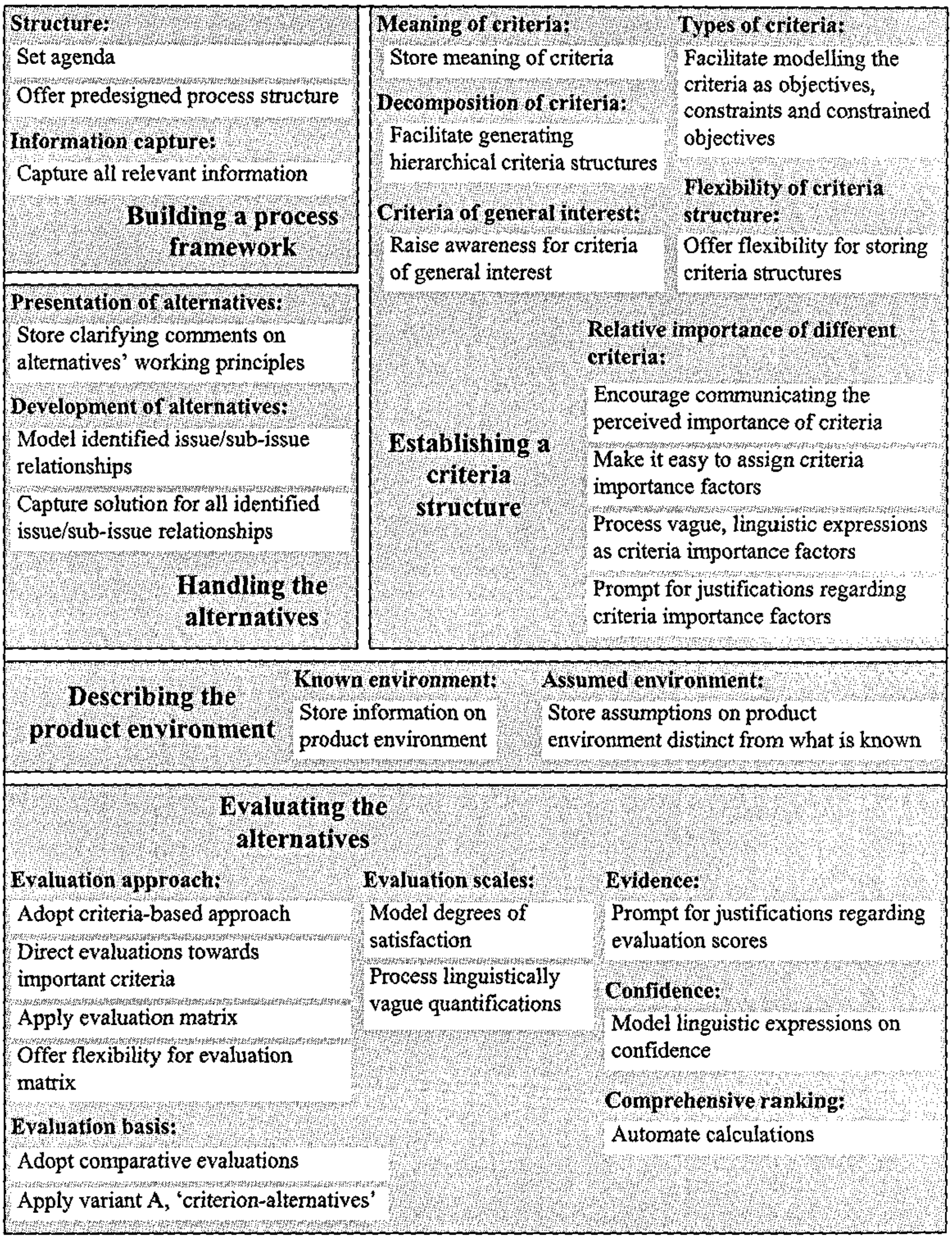


Figure 5.1: Requirements for decision-making support

5.4 Summary

The aim of this chapter was to discuss the implications of the study's results. These implications were expressed as observations on process characteristics and as requirements for decision-making support.

Altogether twelve observations were made. These were related to observations from previous studies. It became evident, that my observations partly support, partly supplement and partly oppose previous observations.

The identified requirements for decision-making support were related to the decision-making methods discussed in chapter 2 and to requirements as suggested from previous studies.

Relating the identified requirements to the discussed decision-making methods revealed that none of them fully satisfies all requirements. According to these requirements effective decision-making support for engineering design involves implementing a flexible, computerised information management system. The system may be based on a decision-making method similar to the Analytic Hierarchy Process (AHP) possibly including fuzzy features. This is because the AHP seemed to satisfy the set of requirements to a higher extent than the other methods. Aspects that were considered unsatisfactory referred in particular to the AHP's *precise* quantification of comparisons and the difficult input procedure. It also needs to be considered whether an AHP-like method could cope with modelling the decision-makers' confidence.
Relating the requirements identified in this research to the requirements identified in previous studies revealed that some of them support previous findings, some of them specify general findings of previous studies and some of them are new, not being mentioned or addressed in the previous studies.

Chapter 6

Conclusions

6.1 Outline of the research

The research study described by this thesis is based on the view that the engineering design process is a series of interrelated operations which are driven by decisions. Accepting this view leads to the realisation that an effective and efficient design process relies upon effective and efficient decision-making. As a consequence, supporting decision-making may be a significant means for achieving design process improvements.

The study aimed at improving the engineering design process by contributing towards the development of context independent decision-making support. The particular interest was selection-type decision-making during the conceptual engineering design phase. The research problem, as formulated in chapter 1 is: How to support selection-type decision-making in conceptual engineering design?

Through a literature review, as described in chapter 2, a preliminary solution to the research problem could be found. That is, a variety of methods from different research disciplines were identified, all of which may be of benefit for supporting the resolution of selection problems in conceptual design. Yet, it was realised that the methods are not used widely in industrial practice despite strong indications that practitioners could indeed benefit from decision-making support. The conclusions from the literature review were: (i), there is a need for more empirical research with the aim to identify requirements for decision-making support and (ii), as a prerequisite, a greater, detailed understanding of relevant decision-making processes had to be generated.

A specific methodology was developed in chapter 3 that would allow for the generation of a detailed understanding of decision-making processes and the identification of requirements for decision-making support. The conclusions were that (i) a descriptive study was needed which leads to the generation of a descriptive decision-making process model, (ii) this model would consist of decision-making related activities and decision-making process steps, (iii) for the empirical identification of these activities and steps an exploratory, qualitative research approach set in the constitutive paradigm was suitable and (iv) the main research methods to be involved were real-time protocols for the data collection and content analysis as well as pattern coding for the data analysis.

The data collection was carried out through three workshops in a laboratory setting involving two groups of final year engineering design students and one group of experienced professionals. The groups were given the assignment to evaluate a number of conceptual design solutions and to decide which one should be selected for further refinement. The workshops were recorded and transcribed for the data analysis. The main results of the data analysis, as described in chapter 4, were a set of decision-making related activity categories, including the quantification of their time consumptions across all three groups, and a detailed decision-making process model showing various process steps as well as different process variants. The implications of the study's results were then expressed in chapter 5 as observations on process characteristics and as requirements for decision-making support.

6.2 Main conclusions

At the outset of this research three main research question were formulated. These questions, associated with a number of particular research objectives, were stated in section 1.2. In the following sections the project's main conclusions will be summarised according to each of these questions and objectives in turn.

6.2.1 Conclusions related to research question 1

Research question 1 was formulated as: what are the specific characteristics of existing decision-making support methods? It involved two objectives, 1.1 'identify typical decision-making support methods' and 1.2 'discuss the methods with the aim of understanding their characteristics'.

Objective 1.1: identify typical decision-making support methods.

Objective 1.1 was addressed in chapter 2, in particular in sections 2.2 and 2.3. Typical methods for supporting selection-type decision-making problems as occurring in conceptual design are:

- Multi-attribute Value Theory (MAVT);
- Analytic Hierarchy Process (AHP);
- Outranking methods;
  - Electre;
  - Promethee;
  - Decision-Making with Quality Function Deployment (QFD);
  - Method of controlled convergence;
- Multi-attribute Utility Theory (MAUT);
- Engineering Decision Support System;
- Yang and Sen's Evidential Reasoning Method;
- Fuzzy methods.

Objective 1.2: discuss the methods with the aim of understanding their characteristics.

Objective 1.2 was addressed in chapter 2, in particular in section 2.4. Typical characteristics that allow for distinguishing the methods refer to:

- Evaluation basis: absolute or relative;
- Overall result: rating or ranking;
- Sensitivity (for absolute evaluations): non-boolean or boolean;
- Sensitivity (for comparisons): quantitative or qualitative;
- Indication of performance differences (for rankings): cardinal or ordinal;
- Assignment of criteria-importance weights: direct or comparison;
- Capability of modelling uncertainty: none, performance, weights;

- Type of uncertainty being modelled: objective probability, subjective belief (confidence), subjective belief and disbelief, fuzziness;

6.2.2 Conclusions related to research question 2

Research question 2 was formulated as: what are the requirements for decision-making support? It involved two objectives, 2.1 'develop a detailed understanding of decision-making processes' and 2.2 'identify support requirements through the developed understanding'.

Objective 2.1: develop a detailed understanding of decision-making processes

In section 3.3 it was discussed how a detailed understanding of decision-making processes may be generated. The suggested solution was to generate a descriptive model of such processes. To do so three sub-objectives were formulated: 2.1.1 'identify decision-making related activities', 2.1.2 'quantify the identified activities' relative time consumptions' and 2.1.3 'use the identified activities to model process steps'.

Objective 2.1.1: identify decision-making related activities

Objective 2.1.1 was addressed in chapter 4, in particular in sections 4.3.1 and 4.3.2. A set of twelve main-categories for decision-making related activities was identified. These activities were:

- Discussing the process approach;
- Identifying criteria;
- Defining criteria;
- Weighting criteria;
- Clarifying concept working principles;
- Clarifying the product environment;
- Deliberating sub-issues;
- Gaining external information;
- Raising evidence;
- Determining or evaluating performances;
- Mapping intuition onto ranking;
- Controlling the process.

A number of these activity categories could be sub-divided into various more specific sub-categories.

Objective 2.1.2: quantify the identified activities' relative time consumptions

Objective 2.1.2 was addressed in chapter 4, in particular in section 4.3.3. For each activity category the time consumption was measured in each group's decision-making process. For comparability the time consumptions were expressed in relative terms and then graphically presented in the form of charts as in figure 4.4. It became apparent that the quantity (number of category) / value (time consumption) distribution in the three charts of figure 4.4, which depicts the main-categories, approximate to Pareto distributions. These distributions showed, among other issues, that there were five dominating activity categories across all three group's decision-making processes, regardless of the chosen approach and the group members' professional status. These were:

- Discussing the process approach;
- Deliberating sub-issues;
- Raising evidence;
- Determining or evaluating performances;
- Controlling the process.

Another dominating activity category for the two formal groups was 'defining criteria'. For the informal group another dominating activity category was 'gaining external information'.

Objective 2.1.3: use the identified activities to model process steps

Objective 2.1.3 was addressed in chapter 4, in particular in section 4.4. A decision-making process model was established that incorporates various steps and sub-steps. The model is shown in figure 4.22. There were three main steps observed for all groups: step 1 'prepare the process', step 2 'explore criteria and alternatives' and step 3 'conclude the process'. With respect to step 2 'explore criteria and alternatives' there were two variants of approaches: comprehensive approach and criteria-based approach. The former approach was informal and was characterised by little consistency regarding evaluation criteria. It was applied by the informal students. The latter approach was formal and was characterised by a high level of structure and consistency regarding evaluation criteria. Yet, more time was consumed by this approach. It was applied by the two groups that both used formal evaluation methods.

The different process steps consisted of activities that the groups carried out. For some of these steps particular patterns of activities were modelled.

Observations:
While quantifying the various activity category's time consumptions and while developing the decision-making process model a number of observations were made regarding particular characteristics of the three groups' decision-making processes. These were addressed in chapter 5, in particular in section 5.2. Summarised the main observations can be expressed as:

- The effects of using a formal method seemed to override possible effects of different professional status. An exception was that the professional engineers spent more time than the students did for gathering evidence to justify their evaluations.
- Using a formal, methodological approach led to a higher level of consistency and to more efforts towards understanding the evaluation criteria than observed in the informal approach.
- Generally, the decision-makers were only consistent with respect to those aspects of the decision-making process that were documented.
- Only those decision-makers who used a formal, methodological approach documented aspects of their decision-making processes. However, the extent of documentation and therefore the amount of recorded information was low.
- The aspect of criteria importance was either not considered at all or treated as little more than a formality.
- Utility functions that would have enabled the decision-makers to evaluate the alternatives on an absolute basis did not seem to exist. Instead, the alternatives' performances were compared with each other and then ranked. The tendency to compare, i.e. to rank rather than rate, was particularly evident during the formal processes.
- There were two information formats that could not be expressed by the scales of the evaluation methods used by the formal decision-makers. These formats were linguistically vague quantifications and expressions of confidence in evaluations.

Objective 2.2: identify support requirements through the developed understanding

Objective 2.2 was addressed in chapter 5, in particular in section 5.3. Five major aspects of support were identified. For each of these aspects there are a number of specific requirements. For each aspect in turn they are:

Aspect 1 - Building a process framework; decision-making support should:

- encourage applying formal processes that provide an effective structure and set a guiding agenda;
- offer pre-designed, formal structures for general decision-making processes that may be readily adopted by the decision-makers;
- be capable of capturing all relevant information that is generated during decision-making processes.

Aspect 2 - Establishing a criteria structure; decision-making support should:

- be capable of storing the meaning of criteria, easily accessible at all times;
- facilitate generating and storing hierarchically decomposed criteria structures;
- raise awareness for criteria of general interest;
- facilitate modelling the criteria as objectives, constraints and constrained objectives;
- offer more flexibility for storing and presenting criteria structures than paper-based decompositions;
- encourage communicating the perceived importance of the different evaluation criteria;
- make it easy and straightforward for the decision-makers to express their values as criteria importance factors (weightings);
- be capable of processing vague, linguistic expressions as a valid format for criteria-importance factors;
- include a feature that prompts the decision-makers to associate their criteria-importance factors with an effective justification.

Aspect 3 - Handling the alternatives; decision-making support should:

- facilitate capture of clarifying comments on the alternatives' working principles;
- model identified issue/sub-issue relationships;

- capture an associated solution for identified issue/sub-issue relationships.

Aspect 4 - Describing the product environment; decision-making support should:

- store any information on the product environment generated during the decision-making process;
- store assumptions on the product environment, but keep them distinct from what is known about the environment.

Aspect 5 - Evaluating the alternative solutions; decision-making support should:

- be built upon a criteria-based evaluation approach;
- facilitate focusing the evaluations towards the important criteria;
- apply evaluation matrices;
- offer more flexibility for evaluation matrices than can be achieved by using paper;
- be based on comparative evaluations, but decision-makers should be prompted to mark problem areas (criteria);
- be based on variant A, 'criterion-alternatives' for restricted evaluations;
- model various degrees of satisfaction regarding performance evaluations;
- be capable of processing linguistically vague quantifications for expressing any extent of satisfaction differences between the alternatives' performances;
- include a feature that prompts the decision-makers to associate their evaluation scores with justifying evidence;
- be capable of modelling the decision-makers' confidence in evaluations - it should allow for linguistic terms rather than numeric expressions;
- offer a feature for automating any calculations necessary for applying aggregation functions.

6.2.3 Conclusions related to research question 3

Research question 3 was formulated as: do the identified methods' characteristics match the identified requirements? It involved the objective 3.1 'compare the requirements with the methods' characteristics'.

Objective 3.1: compare the requirements with the methods' characteristics

Objective 3.1 was addressed in chapter 5, in particular in section 5.3. The discussion in chapter 5 suggested that none of the methods as discussed in chapter 2 fully satisfies all identified requirements. According to these requirements effective decision-making support for engineering design involves implementing a flexible, computerised information management system. The system may be based on a decision-making method similar to the Analytic Hierarchy Process (AHP) possibly including fuzzy features. This is because the AHP seemed to satisfy the set of requirements to a higher extent than the other methods. Aspects that were considered unsatisfactory referred in particular to the AHP's *precise* quantification of comparisons and the difficult input procedure. It also needs to be considered whether an AHP-like method could cope with modelling the decision-makers' confidence.

6.3 Contributions

The outcomes of this research project may be seen as contribution towards four different aspects of work in the field: (i) for directing novice researchers, (ii) for theory generation, (iii) for guiding the practitioners and (iv) for the development of decision support systems.

For directing novice researchers

As indicated in chapters 1 and 2, the overall field of decision-making support is vast and very versified. It is therefore especially helpful for novices to start new research projects by gaining an overview of the entire field. Such overview (i) helps appreciating particular research efforts in a larger context, (ii) provides an orientation for further research, (iii) suggests delimitations and (iv) indicates which research disciplines are involved.

This project has contributed towards gaining an overview over the field by providing structured discussions on the aspects involved in the general field of decision-making support (see figure 1.1 and section 1.3.3) and on the disciplines involved in the specific field of design selection support (see figure 2.1 and section 2.1.3).

For theory generation

In chapter 3, section 3.2 the 'greater research process' was discussed and this research project's position within the process was described. It was indicated that this research falls in the realm of descriptive studies, which aim eventually at the generation of a descriptive theory. Yet, as also mentioned in section 3.2 a descriptive theory is not normally an individual effort, but the joint achievement of an entire research community.

This project has contributed towards theory generation by developing a descriptive decision-making process model. In particular, this is the identification of a set of decision-making related activities, as described in chapter 4, section 4.3 and the quantification of their time consumptions as well as the detailed analysis of relationships between individual activities and their representation as a decision-making process model as described in chapter 4, section 4.4.

The research approach of this project differs from approaches taken by previous projects. This was discussed in sections 2.4, 4.2.4 and 4.5. Because of this, the project may be seen as a *novel* research undertaking, and as such, its results are *new* contributions to the body of knowledge in the field of decision-making support in engineering design.

For guiding the practitioners

In section 3.2 it was discussed that, generally speaking, the outcomes of individual descriptive studies may also be used in industry as a basis for guidelines or examples as well as for suggesting aspects that are in need of support.

This project's contributions for practitioners may be seen in particular in the observations as raised and discussed in section 5.2. These may be of immediate interest as they (i) clearly indicate possible consequences from for example adopting informal approaches as opposed to formalised approaches towards decision-making and (ii) identify process shortcomings that could be avoided even without the help of sophisticated support systems. The discussion in section 5.2 also highlighted that these observations partly agree and partly disagree with observations gained from previous studies. Hence, these observations may be seen as a *new* contribution.

For the development of decision-making support systems

Also of rather practical interest may be the identified requirements for decision-making support as raised and discussed in section 5.3. These requirements are applicable for the future development of decision-making support systems.

Relating the requirements identified in this research to the requirements identified in previous studies revealed that some of them support previous findings, some of them specify general findings of previous studies and some of them are new, not being mentioned or addressed in the previous studies. Hence, the set of requirements identified in this research may also be seen as a *new* contribution.

6.4 Limitations

As any research, this study has a number of limitations. These are related to the gathered data, but also to the data analysis. With respect to the gathered data, limitations result from (i) the workshop environment, (ii) the workshop participants, (iii) the task tackled in the workshops and (iv) the chosen means of protocol generation. With respect to the data analysis limitations result from subjectivity.

Limitations are inevitable in research and as such do not generally question the study's value. But one needs to be aware of their origin for a correct interpretation of the study's results and implications.

Workshop environment

The general research problem addressed by this study is, how to support selection-type decision-making in conceptual engineering design? One of the main objectives related to this research problem was to develop a detailed understanding of actual decision-making processes. Actual decision-making processes of such a type take place in the design departments of industrial companies. Yet, this study has gathered its data under laboratory conditions, i.e. in university-based workshops rather than in industry.

It may be expected that laboratory environments affect the gathered data, at least to some extent. On the other hand, as discussed in section 3.4.4, many previous studies also gathered their data under laboratory conditions: the advantage is practicality, in particular the possibility for repeated studies under similar conditions.

Workshop participants

As the chosen workshop environment, the workshop participants did not entirely reflect 'real' conditions either. This is due to the participation of students. Yet, as these students were finishing their final year of a degree-course in engineering design they represented *almost*-fully qualified engineering designers, with little (they had done industrial placements) industrial experience though. As discussed in section 3.4.6, students were 'used' as participants in previous studies: the advantage is practicality, in particular the students' availability and readiness to participate.

Workshop task

The workshop participants were asked to select a solution for a particular design problem. As discussed in section 3.5.1, this design problem was considered suitable for the study as it was based on a 'real' industrial design issue and it had the right level of complexity, i.e. low complexity. This is rather common for studies of this type, but one may still wonder to what extent this reflects industrial conditions.

Protocols

The method of data collection applied for this research was real-time protocols. These are detailed records of the participants' behaviour, especially their *utterances*. As discussed in section 3.4.4, the strength of this method is its ability to record highly detailed data reflecting actual rather than perceived behaviour. However, the weakness is its inability to record *thoughts*. As such, very relevant data, in form of the participants' thoughts, is ignored and therefore lost for understanding and modelling the process.

Subjectivity

To a large extent this research is concerned with generating a descriptive decision-making process model. This model is based on decision-making related activities and on a number of constructs containing activities. These activities and constructs were basically identified and defined by me, the researcher, through qualitative, exploratory research. This means quite clearly that there is a substantial amount of subjectivity involved, even though the activities were compared with an 'established norm' as discussed in section 4.3.2. Obviously, exactly the same therefore applies to the generated process model, to the observations and to the set of requirements for decision-making support.

The issue of subjectivity was discussed in section 3.4.3 and acknowledged by setting the research in a suitable paradigm: the constitutive research paradigm. From a reductionist's perspective, lack of generalisability is the major drawback of research set in the constitutive paradigm. By choosing this paradigm, generalisability was traded-off for 'depth of study' and generation of 'understanding'.

6.5 Further work

Different leads may be followed in efforts taking this work to further stages. Some of them might be directed towards achieving short term goals whereas others might be directed towards achieving long term goals.

6.5.1 Short term

Short terms goals may include the validation of this study's implications and the implementation of a decision support system as prototype.

Validation of this study's implications

This study's implications were expressed in chapter 5 as observations and as requirements for decision-making support. These observations and requirements are mainly based on the three workshops that I analysed. Apart from this, previous, similar studies have been discussed in this respect as well. However, relatively little input has come from industrial sources.

The observations and suggested requirements for decision-making support are now listed and discussed quite clearly and in little abstract terms. Therefore, it appears feasible to use them directly as foundation upon which to conduct semi-structured interviews across a few design departments of industrial companies. Such interviews may on the one hand highlight areas of particular significance. These might be observations that address process characteristics which are readily recognised by practitioners and requirements which they see as especially relevant. On the other hand, such interviews give practitioners the chance to add their own observations and express their own requirements for decision-making support.

The outcomes of the semi-structured interviews may then be used for a questionnaire survey. This survey could be aimed at (i) gaining wider recognition of the identified requirements and (ii) generating an order of relevance.

Decision support system

For the implementation of a decision support system it would be necessary to first design an abstract framework showing how such a system may satisfy the identified requirements. This means that the system features need to be determined.

In a next step the system features have to be translated into a precise software architecture which may then be implemented as a prototype system. The purpose of such a system is (i) to allow for experiments, i.e. comparing system supported decision-making processes with unsupported processes in an experimental study, and (ii) actually provide practitioners with a practical tool for supporting their decision-making processes.

6.5.2 Long term

Long term goals may be associated with the general research process as shown in figure 3.1. Basically, this means that first a descriptive theory and then a prescriptive model need to be established.

Descriptive theory

For the establishment of a descriptive theory it is necessary to accumulate the results of various individual descriptive studies and generalise them. In particular, this may be achieved by repeating this study under different conditions regarding:

- Environment of study;
- Participants of study;
- Tasks / design issues used in study;
- Data collection methods applied in study;
- Data analysis methods applied in study;

These conditions reflect the particular aspects discussed in connection with this study's limitations in section 6.3. As such, a repetition of the study under different conditions would help alleviating these limitations.

Apart from this study's repetition, studies that focus on different and also more specific aspects of selection-type decision-making are needed for the generation of a descriptive theory. A specific aspect that seems to have caused particular difficulties for this study's participants is the prioritisation of criteria. Hence, it would be worthwhile to study this aspect in more detail. Individual descriptive studies suggest hypotheses that may then be tested by quantitative research eventually leading to the generation of a descriptive theory.

Prescriptive model

From the descriptive theory and by also using assumptions and experience, preferably from various fields including the very related psychology disciplines (see figure 1.1), a prescriptive model showing a desired state may be generated. The transition from the descriptive as-is state to the prescriptive as-desired state may then be attempted through the implementation of a decision support method / tool. This method or tool is then again validated and tested in a research cycle as shown in figure 3.1.

References

Ahn, H. and Dyckhoff, H., 1997, 'Flexible Concept for Decision Support: Evaluation and Selection of Suitable Product Development Activities', *European Journal of Operational Research*, Vol.100, No.2, pp.266-276.

Allen, J.K., 1996, 'Decision to Introduce New Technology: The Fuzzy Preliminary Selection Decision Support Problem', *Engineering Optimisation*, Vol.26, No.1, pp.61-77.

Andreasen, M.M. and Olesen, J., 1990, 'The Concept of Dispositions', *Journal of Engineering Design*, Vol.1, No.1, pp.17-36.

Antonsson, E.K. and Otto, K.N., 1995, 'Imprecision in Engineering Design', *Journal of Mechanical Design*, Vol.117B, pp.25-32.

Argyris, C., 1985, *'Strategy, Change and Defensive Routines'*, Pitman publishers.

Azani, H. and Khorramshahgol, R., 1990, 'Analytic Delphi Method (ADM): A Strategic Decision Making Model Applied to Location Planning', *Engineering Costs and Production Economics*, Vol. 20, pp.23-28.

Barton, J.A, Love, D.M. and Taylor, G.D., 2000, 'Design Determines 70 Percent of Cost? A Review of Implications for Design Evaluation', *Proceedings of the Engineering Design Conference, EDC'00*, UK, pp.247-258.

Barzilai, J., 1997, 'A New Methodology for Dealing with Conflicting Engineering Design Criteria', *Proceedings of the 1997 National Conference of the American Society for Engineering Management*, USA, pp.73-79.

Bellmann, R.E. and Zadeh, L.A., 1970, 'Decision-making in a Fuzzy Environment', *Management Science*, Vol.17, No.4, pp.141-164.

Berg, B.L., 1989, *'Qualitative Research Methods for the Social Sciences'*, Allyn and Bacon publishers, USA.

Blessing, L., Chakrabarti, A. and Wallace, K., 1995, 'A Design Research Methodology', *Proceedings of the International Conference on Engineering Design, ICED'95*, Czech Republic, pp.50-55.

Bradburn, B.G. and Stauffer, L.A., 1991, 'A Comparison of Knowledge Elicitation Methods', *Proceedings of the International Conference on Engineering Design, ICED'91*, Switzerland, pp.298-305.

Brans, J.P., Vincke, Ph. and Mareschal, B., 1986, 'How to Select and How to Rank Projects: The Promethee Method', *European Journal of Operational Research*, Vol.24, pp.228-238.

Bryman, A., 1988, *'Quality and Quantity in Social Research'*, Unwin Hyman publishers.

Bryson, N. and Mobolurin, A., 1994, 'An Approach to Using the Analytic Hierarchy Process for Solving Multiple Criteria Decision Making Problems', *European Journal of Operational Research*, Vol. 76, pp.440-454.

Buchanan, J.T., 1994, 'An Experimental Evaluation of Interactive MCDM Methods and the Decision Making Process', *Journal of the Operational Research Society*, The Operational Research Society, UK, Vol. 45, No. 9, pp.1050-1059.

Cannon, D.M. and Leifer, L.J., 1999, 'The Place of Design Research in the World of Contemporary Philosophical Thought, and Vice-Versa: An Illustrative Example', *Proceedings of the International Conference on Engineering Design, ICED'99*, Germany, Vol. 3, pp.1635-1640.

Carnahan, J.V., Thurston, D.L. and Liu, T., 1994, 'Fuzzy Ratings for Multiattribute Design Decision Making', *Journal of Mechanical Design*, Vol.116, No.2, pp.511-521.

Chen, J.L. and Lee, C.-R., 1993, 'Qualitative Programming Method for Qualitative Design Evaluation in the Conceptual Design Stage', *Journal of Engineering Design*, Vol.4, No.3, pp.213-220.

Choi, H., Suh, E. and Suh, C., 1994, 'Analytic Hierarchy Process: It Can Work for Group Decision Support Systems', *Computers & Industrial Engineering*, Vol.27, No.1-4, pp.167-171.

Churchill, Jr. G.A., 1995, *'Marketing Research-Methodology Foundations'*, Dryden Press publishers.

Colton, J.S. and Pun, R.C., 1994, 'Information Frameworks for Conceptual Engineering Design', *Engineering with Computers*, pp.22-32.

Conrath, D.W., 1973, 'From Statistical Decision Theory to Practice: Some Problems with the Transition', *Management Science*, Institute of Management Sciences, USA, Vol.19, pp.873-883.

Court, A.W., 1998, 'Integrating Information and Knowledge in New Product Development: A View on the Pertinent Research Issues', *Proceedings of the Engineering Design Conference, EDC'98*, UK, pp.495-502.

Cross, N., 1994, *'Engineering Design Methods: Strategies for Product Design'*, Chichester Wiley publishers.

Cryer, P., 1999, *'The Research Student's Guide To Success'*, Open University Press.

Cziulik, C. and Driscoll, J., 1997, 'Conceptual Design Evaluation Based on Function Metrics', *Proceedings of the International Conference on Engineering Design, ICED'97*, Finland, pp.411-416.

Denzin, N.K. and Lincoln, Y.S., 1994, 'Introduction: Entering the Field of Qualitative Research' in: *The Handbook of Qualitative Research* by Denzin, N.K. and Lincoln, Y.S., Sage Publications, pp.1-17.

Diakoulaki, D. and Koumoutsos, N., 1991, 'Cardinal Ranking of Alternative Actions: Extension of the Promethee Method', *European Journal of Operational Research*, Vol.53, pp.337-347.

Dixon, J.R., 1988, 'On Research Methodology Towards a Scientific Theory of Engineering Design', *Journal of Artificial Intelligence in Engineering Design, Analysis, and Manufacturing (AI EDAM)*, Vol.1, No.3, pp.316-337.

Dubois, D. and Prade, H., 1978, 'Operations on Fuzzy Numbers', *International Journal of Systems Sciences*, Vol.9, No.6, pp.613-626.

Duffy, A.H.B., Duffy, S.M., Andreasen, M.M. and Rimmer, D., 1995, 'Team Engineering within the Context of Product Design Development', *Report*, The CAD Centre, DMEM, University of Strathclyde, Glasgow, UK.

Dwarakanath, S., 1996, *'Understanding and Supporting Decision Making in Engineering Design'*, PhD. Thesis, Cambridge University, UK.

Dwarakanath, S. and Wallace, K.M., 1995, 'Decision-making in Engineering Design: Observations from Design Experiments', *Journal of Engineering Design*, Vol.6, No.3, pp.191-206.

Easterby-Smith, M. , Thorpe, R. and Lowe, A., 1995, *'Management Research: An Introduction'*, Sage Publications.

Efstathiou, J., 1984, 'Practical Multi-Attribute Decision-Making and Fuzzy Set Theory', *Fuzzy Sets and Decision Analysis,* TIMS/Studies in Management Science, Vol.20, pp.307-320.

Ehrlenspiel, K. and Dylla, N., 1993 'Experimental Investigation of Designers' Thinking Methods and Design Procedures', *Journal of Engineering Design*, Vol.4, No.3, pp.201-211.

Ehrlenspiel, K. and Lenk, E., 1993, 'Einflüsse auf den Bewertungsprozeß beim Konstruieren', *Proceedings of the International Conference on Engineering Design, ICED'93*, The Netherlands, pp.449-456.

Eide, P.G., 1997, 'Research Design in Design Research', *Proceedings of the International Conference on Engineering Design, ICED'97*, Finland, pp.377-380.

Emory, C.W. and Cooper, D.R., 1991, *'Business Research Methods'*, Irwin Homewood publishers.

Feldy, E.C., 1997, 'Introduction to Decision Making in Design Engineering', *Proceedings of the National Design Engineering Conference at National Manufacturing Week*, USA, pp.153-158.

Franceschini, F. and Rossetto, S., 1995, 'QFD: The Problem of Comparing Technical/Engineering Design Requirements', *Research in Engineering Design*, Vol.7, No.4, pp.270-278.

French, M.J., 1985, *'Conceptual Design for Engineers'*, Springer-Verlag.

Geng, H., Bode, J., Zhang, X. and Shouju, R., 1996, 'CODESS – A System for Decision Support in the Conceptual Phase of New Product Development', *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, USA, Vol.2, pp.1388-1392.

Gero, J.S. and McNeill, T., 1998, 'An Approach to the Analysis of Design Protocols', *Design Studies*, Vol.19, No.1, pp.21-61.

Ghotb, F. and Warren, L., 1995 'Case Study Comparison of the Analytic Hierarchy Process and a Fuzzy Decision Methodology', *Engineering Economist*, Vol.40, No.3. pp.233-246.

Graham, I. and Jones, P.L., 1988, *'Expert Systems - Knowledge, Uncertainty and Decision'*, Chapman and Hall publishers.

Guba, E.G. and Lincoln, Y.S., 1989, *'Fourth Generation Evaluation'*, Sage Publications.

Hales, R.F., 1995, 'Quality Function Deployment as a Decision Making Tool', *Proceedings of the International Conference of the American Production and Inventory Control Society*, USA, pp.489-492.

Haque, B. and Pawar, K.S., 1998, 'Development of a Methodology and Tool for Analysing New Product Design and Development in a Concurrent Engineering Environment', *Proceedings of the Engineering Design Conference, EDC'98*, UK, pp.669-678.

Hartvig, S.C., 1998, 'Software for Evaluation of Conceptual Design', *Proceedings of the International Computing Congress in Civil Engineering*, USA, pp.811-821.

Hauser, J. and Clausing, D., 1988, 'The House of Quality', *Harvard Business Review*, pp.73-73.

Hazelrigg, G.A., 1997, 'On Irrationality in Engineering Design', *Journal of Mechanical Design*, Vol.119, No.2, pp.194-196.

Hazelrigg, G.A., 1998, 'A Framework for Decision-Based Engineering Design', *Journal of Mechanical Design*, Vol.120, pp.653-658.

Herling, D., 1997, *'An Engineering Decision Support System (EDSS) with Alternative-Criterion Pair Evaluations'*, PhD thesis, Oregon State University, USA.

Herling, D., Ullman, D.G. and D'Ambrosio, B., 1995, 'Engineering Decision Support System (EDSS)', *Proceedings of the Design Engineering Technical Conferences, DETC'95*, USA, Vol.83, No.2, Pt.1, pp.619-626.

Höhne, G., 1997, 'Design Process Optimisation by Decision-Making During the Conceptual Design Phase', *Proceedings of the International Conference on Engineering Design, ICED'97*, Finland, pp.169-174.

Hsiao, S.-W., 1998, 'Fuzzy Logic Based Decision Model for Product Design', *International Journal of Industrial Ergonomics*, Vol.21, No.2, pp.103-116.

van Huylenbroeck, G., 1995, 'The Conflict Analysis Method: Bridging the Gap Between ELECTRE, PROMETHEE and ORESTE', *European Journal of Operational Research*, Vol.82, pp.490-502.

Ida, M., Katai, O., Sawaragi, T. and Iwai, S., 1992, 'Decision Making Under Fuzziness with Multiple Objectives and its Application to Design Problem Solving', *Proceedings of the International Conference on Fuzzy Logic and Neural Networks*, Japan, Vol.1, pp.325-328.

Ivezic, N. and Garrett, J.H. Jr., 1994, 'Machine Learning Decision Support System for Collaborative Design', *Journal of Computing in Civil Engineering*, No.1, pp.901-908.

Jain, R., 1976, 'Decision Making in the Presence of Fuzzy Variables', *IEEE Transactions on Systems, Man and Cybernetics*, Vol.SMC-6, No.10, pp.698-703.

Jebb, A. and Woolliams, P., 2000, 'Synergizing Engineering Design with Business Performance – An Integrated Methodology', *Proceedings of the Engineering Design Conference, EDC'00*, UK, pp.279-290.

Jerčić, I. and Bajić, D., 1993, 'Analytic Decision-Making Process in Product Design', *Proceedings of the International Conference on Advanced Manufacturing Systems and Technology, AMST'93*, Italy, pp.151-158.

Kaplan, R.S., 1986, 'The Role for Empirical Research in Management Accounting', Accounting, *Organisations and Society*, Vol.11, No.7, pp.15-26.

Keeney, R.L. and Raiffa, H., 1976, *'Decisions with Multiple Objectives: Preferences and Value Tradeoffs'*, John Wiley & Sons publishers, UK.

Kennedy, P., 1997, 'Research in Design Using Methods and Ideas from Social Science Research', *Proceedings of the International Conference on Engineering Design, ICED'97*, Finland, pp.449-454.

De Keyser, W. and Peeters, P., 1996, 'A Note on the Use of Promethee Multicriteria Methods', *European Journal of Operational Research*, Vol.89, pp.457-461.

Khorramshahgol, R. and Moustakis, V., 1988, 'Delphic Hierarchy Process (DHP): A Methodology for Priority Setting Derived from the Delphi Method and Analytical Hierarchy Process', *European Journal of Operational Research*,Vol.37, pp.347-354.

Lacksonen, T., 1995, 'Statistical Decision Analysis', *Proceedings of the 4th Industrial Engineering Research Conference*, USA, pp.564-568.

Law, W.S. and Antonsson, E.K., 1995, 'Hierarchical Imprecise Design with Weights', *Proceedings of the 1995 IEEE International Conference on Fuzzy Systems*, Japan, Part 1, pp.383-388.

Leedy, P., 1989, *'Practical Research'*, Macmillan publishers.

Libertore, M.J. and Stylianou, A.C., 1995, 'Expert Support Systems for New Product Development Decision Making: A Modelling Framework and Applications', *Management Science*, Institute of Management Sciences, USA, Vol.41, No.8, pp.1296-1316.

Lincoln, Y.S. and Guba E.G., 1985, *'Naturalistic Enquiry'*, Sage Publications.

Maffin, D., 1998, 'Engineering Design Models: Context, Theory and Practice', *Journal of Engineering Design*, Vol.9, No.4, pp.315-327.

Marginson, D.E.W., 1998, 'The Case Study, the Researcher and the Issues: Forewarned is Forearmed', *Report*, The Business School, Loughborough University, UK.

Marples, D.L., 1960, 'The Decisions of Engineering Design', *Report*, The Institution of Engineering Designers (IED), UK.

Matthews, P.C., Blessing, L.T.M. and Wallace, K.M., 1999, 'Conceptual Evaluation Using Neural Networks', *Proceedings of the International Conference on Engineering Design, ICED'99*, Germany, 1777-1780.

Midland, T., 1997, 'A Decision Directed Design Approach', *Engineering Designer*, The Institution of Engineering Designers, UK, May/June, pp.4-7.

Miles, M.B. and Huberman, A.M., 1994, *'Qualitative Data Analysis: an Expanded Sourcebook'*, Sage Publications.

Mizrahi, A. and Sullivan, M., 1993, *'Mathematics for Business, Life Sciences, and Social Sciences'*, John Wiley & Sons publishers, UK.

Müller, K. and Sebastian, H.-J., 1997, 'Intelligent Systems for Engineering Design and Configuration Problems', *European Journal of Operational Research*, Vol.100, pp.315-326.

Mutchnik, R.J., Berg, B.L., 1995, *'Research Methods for the Social Sciences - Practice and Applications'*, Allyin and Bacon publishers.

Naudé, P., Lockett, G. and Holmes, K., 1997, 'A Case Study of Strategic Engineering Decision Making Using Judgmental Modelling and Psychological Profiling', *IEEE Transactions on Engineering Management*, USA, Vol.44, No.3, pp.237-247.

Nichols, K., 1992, 'Better, Cheaper, Faster Products - by Design', *Journal of Engineering Design*, UK, Vol.3, No.3, pp.217-228.

Nidamarthi, S., 1999, *'Understanding and Supporting Requirement Satisfaction in the Design Process'*, PhD. Thesis, Cambridge University, UK.

Nidamarthi, S., Chakrabarti, A. and Bligh, T.P., 1999, 'Techniques for Research Method Validation in Protocol Studies', *Proceedings of the International Conference on Engineering Design, ICED'99*, Germany, pp.1571-1576.

Nojiri, H., 1982, 'A Model of the Executive's Decision Processes in New Product Development', *Fuzzy Sets and Systems*, Vol.7, pp.227-241.

Otto, K.N. and Antonsson, E.K., 1991, 'Trade-off Strategies in Engineering Design', *Research in Engineering Design*, Vol.3, No.2, pp.87-104,

Pahl, G. and Beitz, W., 1988, *'Engineering Design – a Systematic Approach'*, The Design Council, UK.

Patterson, D.W., 1990, *'Introduction to Artificial Intelligence & Expert Systems'*, Prentice Hall publishers.

Perry, C., 1996, 'How to Write a Doctoral Thesis – PhD / DPhil', *Paper presented to the ANZ Doctoral Consortium*, University of Sydney, Australia, 1994, with later additions from 17 July 1996, Brunel University, UK.

Platt, J., 1988, 'What Can Case Studies Do?', in: Burges, R.G. (Ed.) *'Studies in Qualitative Methodology: Conducting Qualitative Research'*, JAI Press, UK.

Pomerol, J.-Ch., 1997, 'Artificial Intelligence and Human Decision Making', *European Journal of Operational Research*, Vol.99, pp.3-25.

Pöyhönen, M.A., Hämäläinen, R.P. and Salo, A.A., 1997, 'An Experiment on the Numerical Modelling of Verbal Ratio Statements', *Journal of Multi-Criteria Decision Analysis*, UK, Vol.6, pp.1-10.

Pöyhönen, M. and Hämäläinen, R.P., 2000, 'On the Convergence of Multiattribute Weighting Methods', Helsinki University of Technology, Finland, *Paper accepted for publication in European Journal of Operational Research*

Pugh, S., 1990, *'Total Design'*, Addison-Wesley publishers.

Robson, C., 1993, *'Real World Research'*, Blackwell publishers.

Roozenburg, N.F.M. and Eekels, J., 1995 , *'Product Design: Fundamentals and Methods'*, Jon Wiley & Sons publishers.

Roy, B., 1991, 'The Outranking Approach and the Foundations of Electre Methods', *Theory and Decision*, Vol.31, pp.49-73.

Saaty, T.L., 1990, 'How to Make a Decision: The Analytic Hierarchy Process', *European Journal of Operational Research*, Vol.48, pp.9-26.

Sage, A.P., 1995, 'Risk Management Systems Engineering', *Transactions of the IEEE*, USA, pp.1033-1038.

Salo, A.A. and Hämäläinen, R.P., 1997, 'On the Measurement of Preferences in the Analytic Hierarchy Process', *Journal of Multi-Criteria Decision Analysis*, Vol.6, pp.309-319.

Schlüter, A., 1999a, 'Requirements on the Use of Methods in Industry', *Proceedings of the International Conference on Engineering Design, ICED'99*, Germany, pp.231-232.

Schlüter, A., 1999b, 'The Use of Design Methods in Mid-Size Companies', *Proceedings of the International Conference on Engineering Design, ICED'99*, Germany, pp.405-408.

Schwartz, J., 1962, 'The Pernicious Influence of Mathematics in Science', *Proceedings of the 1960 International Congress on Logic, Methodology and Philosophy of Science*, USA, pp.356-360.

Sen, P., 1995, 'Many Approaches in Evaluation and Decision-Making', *Journal of Engineering Design*, Vol.6, No.3, pp.171-172.

Sen, P. and Yang, J.-B., 1995, 'Multiple-criteria Decision-making in Design Selection and Synthesis', UK, Vol.6, No.3, pp.207-230.

Shah, J.J., 1998, 'Design Research and Industry Relevance', *Proceedings of the Engineering Design Conference, EDC'98*, UK, pp.31-41.

Simpson, L., 1996, 'Do Decision Makers Know What They prefer?: MAVT and ELECTRE II', *Journal of the Operational Research Society*, The Operational Research Society, UK, Vol.47, pp.919-929.

Smith, R.P. and Leong, A., 1998, 'An Observational Study of Design Team Process: A Comparison of Student and Professional Engineers', *Journal of Mechanical Design*, Vol.120, pp.636-642.

Sommer, R. and Sommer, B.B., 1980, *'A Practical Guide to Behavioural Research'*, Oxford University Press, UK.

Stacey, M., Eckert, C., 1999, 'An Ethnographic Methodology for Design Process Analysis', *Proceedings of the International Conference on Engineering Design, ICED'99*, Germany, pp.1565-1570.

Starkey, C.V., 1992, *'Engineering Design Decisions'*, Hodder & Stoughton publishers, UK.

Stauffer, L., 1987, *'An Empirical Study on the Process of Mechanical Design'*, PhD Thesis, Department of Mechanical Engineering, Oregon State University, USA.

Stauffer, L.A., Diteman, M. and Hyde, R., 1991, 'Eliciting and Analysing Subjective Data about Engineering Design', *Journal of Engineering Design*, Vol.2, No.4, pp.351-366.

Tanimoto, S.L., 1990, *'The Elements of Artificial Intelligence'*, Freeman and Company.

Tebay, R., Atherton, J. and Wearne, S.H., 1984, 'Mechanical Engineering Design Decisions: Instances of Practice compared with Theory', *Proceedings of the Institution of Mechanical Engineers*, UK, Vol.198B, No.6, pp.87-96.

Thurston, D.L., 1990, 'Subjective Design Evaluation with Multiple Attributes', *Proceedings of the Design Engineering Technical Conferences, DETC'90*, USA, Vol.18, pp.355-361.

Thurston, D.L. and Carnahan, J.V., 1992, 'Fuzzy Ratings and Utility Analysis in Preliminary Design Evaluation of Multiple Attributes', *Journal of Mechanical Design*, Vol.114, No.4, pp.648-658.

Ullman, D.G., 2000, 'Toward the Ideal Mechanical Engineering Design Support System', Department of Mechanical Engineering, Oregon State University, USA, *paper submitted for publication*, http://www.engr.edu/~ullman.

Ullman, D.G. and D'Ambrosio, B., 1995, 'A Taxonomy for Classifying Engineering Decision Problems and Support Systems', *Proceedings of the Design Engineering Technical Conferences, DETC'95*, USA, Vol.83, No.2, Pt.1, pp.627-637.

Ullman, D.G., Dietterich and T.G, Stauffer, L.A., 1988, 'A Model of the Mechanical Design Process Based on Empirical Data', *Journal of Artificial Intelligence in Engineering Design and Manufacturing (AI EDAM)*, Vol.2, No.1, pp.33-52.

Ullman, D.G., Herling, D., D'Ambrosio, B., 1997, 'What to do Next: Using the Problem Status to Determine the Course of Action', *Research in Engineering Design*, Vol.9, pp.214-227.

Vadde, S., Allen, J.K. and Mistree, F., 1994 'Bayesian Compromise Decision Support Problem for Multilevel Design Involving Uncertainty', *Journal of Mechanical Design*, Vol.116, No.2 , pp.388-395.

Vetschera, R., 1994, 'Estimating Aspiration Levels from Discrete Choices - Computational Techniques and Experiences', *European Journal of Operational Research*, Vol.76, No.3, pp.455-465.

de Vries, T.J.A., van Amerongen, J. and Breedveld, P.C., 1993, 'A Model of the Design Process for Development of CAE Systems', *Proceedings of the International Conference on Engineering Design, ICED'93*, Netherlands, pp.1409-1417.

Wallace, K. and Burgess, S., 1995, 'Methods and Tools for Decision Making in Engineering Design', *Design Studies*, Vol.16, pp.429-446.

Wang, J., 1994, 'Artificial Neural Networks Versus Natural Neural Networks', *Decision Support Systems*, Vol.11, pp.415-429.

Weiss, M.P. and Hari, A., 1997, 'Problems of Concept Selection in Real Industrial Environment', *Proceedings of the International Conference on Engineering Design, ICED'97*, Finland, pp.723-728.

Wolters, W.T.M. and Mareschal, B., 1993, 'Novel Types of Sensitivity Analysis for Additive MCDM Methods', *European Journal of Operational Research*, Vol.81, No.2, pp.281-290.

Wood, K.L. and Antonsson, E.K., 1987, 'A Fuzzy Sets Approach to Computational Tools for Preliminary Engineering Design', *American Society of Mechanical Engineers, Design Engineering Division*, USA, Vol.10, No.1, pp263-171.

Wright, I., 1998, *'Design Methods in Engineering and Product Design'*, McGraw Hill publishers, UK.

Yang, J.-B. and Sen, P. , 1994, 'A General Multi-Level Evaluation Process for Hybrid MADM with Uncertainty', *IEEE Transactions on Systems, Man, and Cybernetics*, USA, Vol.24, pp.1458-1473.

Yang, J.-B. and Sen, P., 1997, 'Multiple Attribute Design Evaluation of Complex Engineering Products Using the Evidential Reasoning Approach', *Journal of Engineering Design*, Vol.8, No.3, pp.211-230.

Yang, J.-B. and Singh M.G., 1994, 'An Evidential Reasoning Approach for Multiple Attribute Decision Making with Uncertainty, *IEEE Transactions on Systems, Man, and Cybernetics*, USA, 24, pp.1-18.

Yao, J.T.P. and Furuta, H., 1986, 'Probabilistic Treatment of Fuzzy Events in Civil Engineering', *Probabilistic Engineering Mechanics*, Vol.1, No.1, pp.58-64.

Yen, J. and Tiao, W.A., 1997, 'Impact Analysis of Design Alternatives on Imprecise Requirements', *Proceedings FUZZ-IEEE'97*, pp.1215-1220.

Yin, R.K., 1994, *'Case study research: Design and methods'*, (Applied Social Research Methods Series, Vol.5), Sage Publications.

Zadeh, L.A., 1965, 'Fuzzy Sets', *Information and Control*, Vol.8, pp.338-353.

Zahedi, F., 1986, 'The Analytic Hierarchy Process - A Survey of the Method and its Applications', *Interfaces,* The Institution of Management Sciences, USA, Vol.16, No.4, pp.96-108.

Žavbi, R., Duhovnik, J., 1996, 'The Analytic Hierarchy Process and Functional Appropriateness of Components of Technical Systems', *Journal of Engineering Design*, Vol.7, No.3, pp.313-327.

Zimmermann, H.J., 1983, 'Using Fuzzy Sets in Operational Research', *European Journal of Operational Research,* Vol.13, pp.201-216.

Zimmermann, H.J., 1996, 'Fuzzy Logic on the Frontiers of Decision Analysis and Expert Systems', *Proceedings of the Conference of the North American Fuzzy Information Processing Society,* USA, pp.65-69.

Zimmermann, H.-J. and Sebastian, H.-J., 1995, 'Intelligent System Design Support by Fuzzy Multi-Criteria Decision Making and/or Evolutionary Algorithms', *Proceedings FUZZ-IEEE'95,* Japan, pp.367-374.

Appendix A

Models of the Design Process

The main text refers to three examples of design process models: French (1985), Pahl and Beitz (1988) and Pugh (1990). These models will be shown graphically and briefly introduced in the following.

French's model

French's model is shown in figure A1. In his 'block diagram' the circles represent stages and the rectangles represent work in progress.

Analysis of the problem aims at identifying the exact *need* to be satisfied. The output is a *statement of the problem* containing three elements:

- A statement of the design problem proper;
- Any limitations placed upon the solution;
- The criterion of excellence.

French believes, the ultimate criterion of excellence is always cheapness, arguing that all other criteria can eventually be reduced to cost and claiming that the best design is the cheapest. Yet, he also states that the ultimate criterion of low cost is often inconveniently remote. Hence, more immediate criteria, such as 'weight', become of greater usefulness.

French points out that from the actual design work there is feedback to the analysis of the

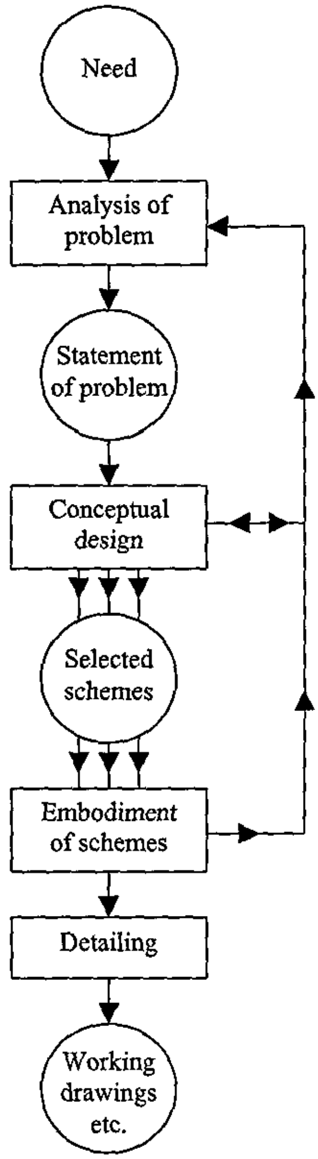


Figure A1: French's block diagram of the design process.

problem. This is because he does not see needs as being absolute, but relative to the costs of fulfilling them.

Conceptual design takes the *statement of the problem* and generates broad solutions in form of 'schemes'. French stresses that conceptual design offers the most scope for striking design improvements and that the most important decisions are made during this phase.

During the *embodiment of schemes* they are worked up in greater detail. If a number of schemes was generated, only few *selected schemes* are worked up and eventually a final selection needs to be made. The output of this phase is a set of general arrangement drawings. French's model indicates there is feedback from this phase to the conceptual design phase, which is why he advocates overlapping the two.

Detailing is the last design phase. A very large number of small but essential points need to be specified. The output is a set of detailed drawings and other documents required for manufacturing the product.

Pahl and Beitz's model

Pahl and Beitz's model maps the flow of work during the design process onto four main phases similar to French's model:

- Clarification of the task;
- Conceptual design;
- Embodiment design;
- Detail design.

Figure A2 shows these phases step by step. Pahl and Beitz point out that after each step, a decision has to be made as to whether the next step should be commenced or whether previous steps have to be repeated.

The first phase, *clarification of the task*, involves collecting information about the requirements and constraints on the solution. Its output is a detailed specification or requirements list.

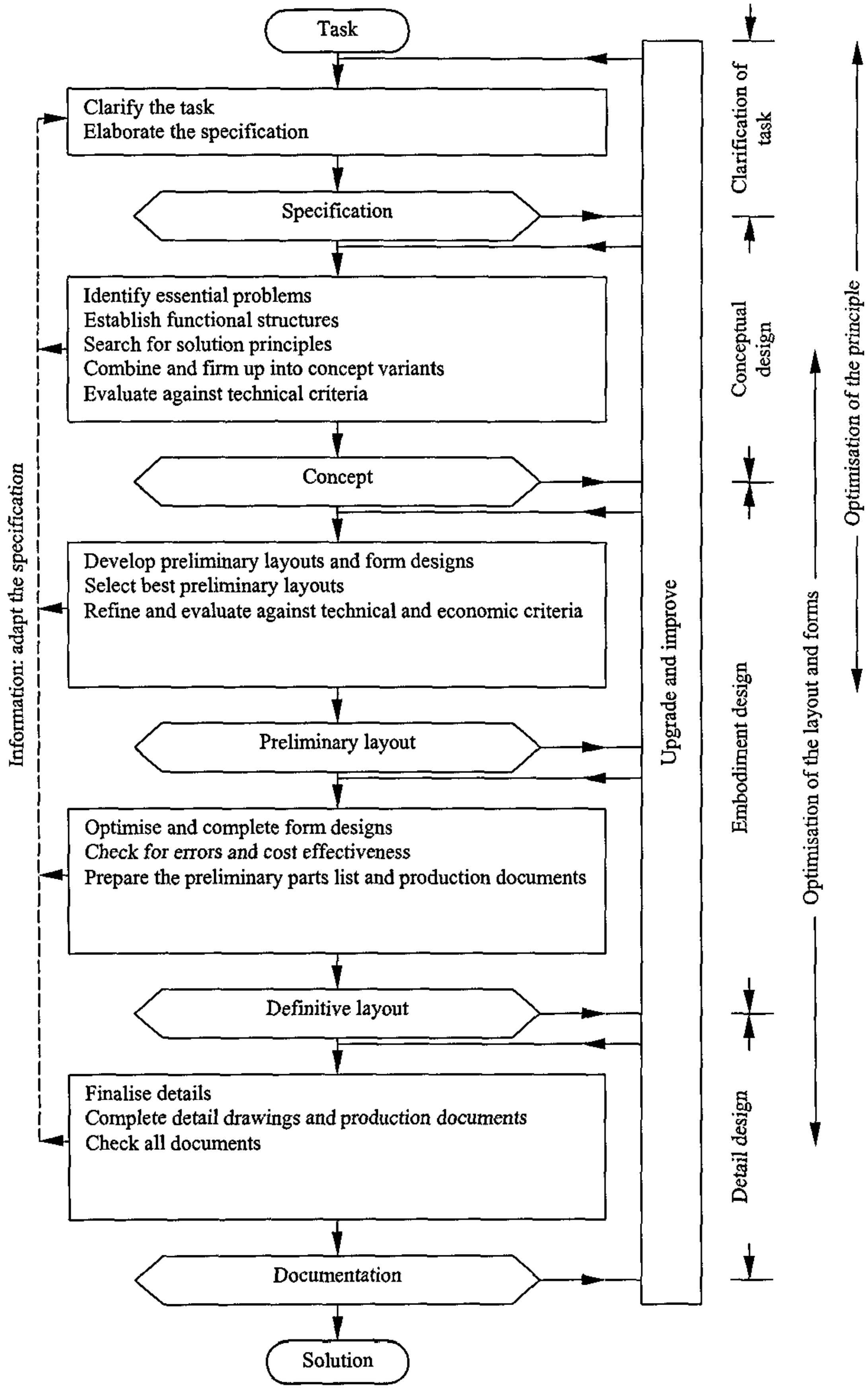


Figure A2: Pahl and Beitz's steps of the design process.

The second phase, *conceptual design*, involves the establishment of function structures, the generation of solutions and their combination into concept variants. To converge at one or few concepts the generated concept variants are then evaluated against the various demand criteria as detailed in the specification or requirements list. Pahl and Beitz suggest, for these evaluations the chief criteria are of a technical nature, though rough economic criteria may also play a part. They stress the particular significance of the conceptual design phase for eventually gaining a successful product when stating that such product is more likely to spring from choosing the most appropriate concept than from exaggerated concentration on details.

The third phase, *embodiment design*, involves determining the layout and forms of the selected concept. To explore the advantages and disadvantages of different approaches, Pahl and Beitz recommend to generate a number of preliminary layouts and then select one that appears particularly promising but which may still benefit from refinement and optimisation on the basis of technical and also economic criteria. The definitive layout provides a check of function, strength, spatial compatibility and it allows for an assessment of financial viability.

The last phase, *detail design*, involves determining the exact arrangements, dimensions, surface properties, etc. of all individual parts. Drawings are generated and the required production documentation is prepared.

Pahl and Beitz see the crucial activities in their model as:

- Optimisation of the principle and
- Optimisation of the layouts and forms.

These two activities influence each other and, as shown in figure A2, overlap to a considerable extent.

Pugh's model

Pugh created a model depicting 'total design'. He defines 'total design' as "*the systematic activity necessary, from the identification of the market/user need, to the selling of the successful product to satisfy that need – an activity that encompasses product, process, people and organisation.*" (Pugh, 1990, p.5)

The model, as shown in figure A3, has a central core of aspects that are imperative for any design, regardless of domain. This core consists of market (user need), product design specification, conceptual design, detail design, manufacture and sales.

Pugh states that all design starts with a need that either fits into a marked or creates a market of its own. From the statement of the need (also called design brief) a product design specification (PDS) must be formulated. The PDS will act as a cloak that envelops all subsequent phases of the central core, thus acting as a control device for total design. As indicated by the vertical double-headed arrows, the main design flow, from market to sales is seen as an iterative process.

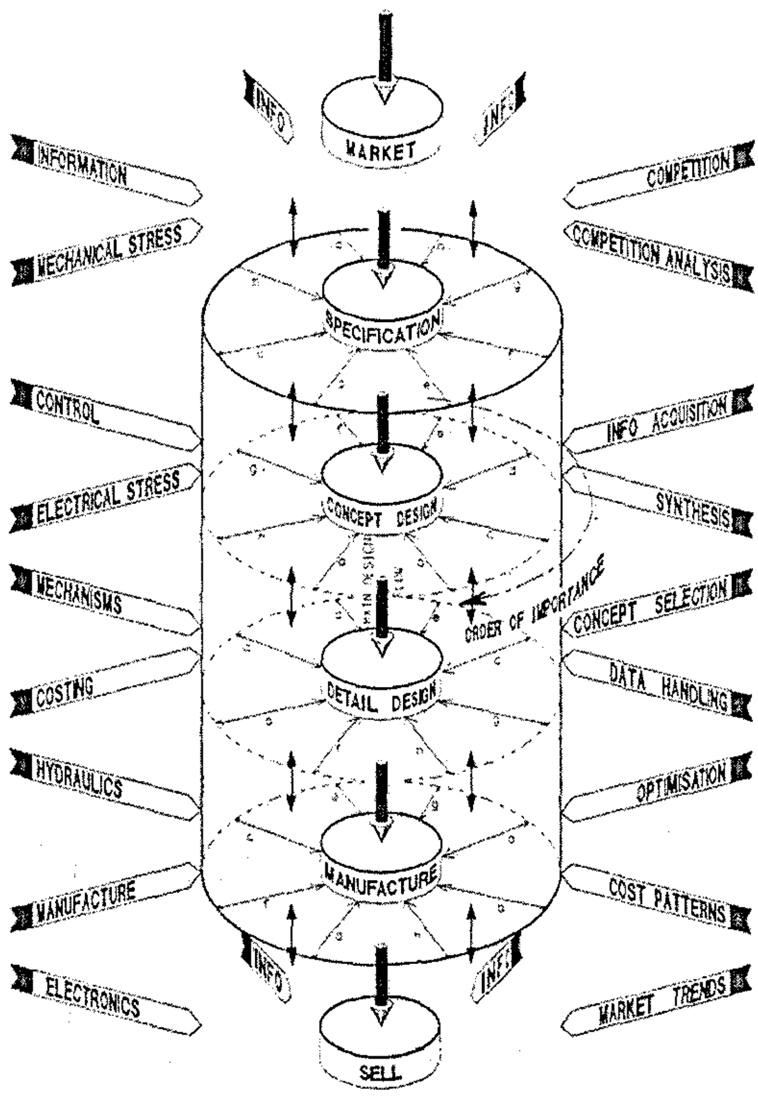


Figure A3: Pugh's model of the design core enveloped by PDS and various inputs.

Pugh's model also shows a number of inputs related to the central core. These refer to two types of a designer's 'tool-kit'. On the right-hand side there are context independent techniques of analysis, synthesis, decision-making, modelling, etc., which could be applied to any product development process. On the left-hand side there are context dependent techniques and technological knowledge, such as stress analysis, thermodynamic analysis, information on materials, electronics, etc. As such, the model depicts a broad design core, enveloped by the PDS, and with inputs from the two types of a designer's tool-kit, i.e. context (discipline/technology) dependent and context independent.

Appendix B

Frameworks of Support Requirements

Frameworks of requirements for decision-making support have been established in particular by two previous studies: Ullman and D'Ambrosio (1995) and Dwarakanath (1996). As discussed in section 2.4.3 these frameworks were established through observations from laboratory based design experiments.

Ullman and D'Ambrosio's study

Ullman and D'Ambrosio developed a decision problem taxonomy, as shown in figure B1. This taxonomy is a scheme for classifying decision-making situations according to a variety of categories. It also allows for classifying existing decision-making methods and tools according to their support regarding these categories. In the same way it specifies requirements for future decision support developments.

Structure	Decision Space	1. Problem Completeness
		2. Abstraction Level
		3. Determinism
	Preference Model	4. Objective Function
		5. Consistency
		6. Comparison Basis
	Belief Model	7. Dimension
		8. Belief Completeness
Focus		9. Problem Focus
Range		10. Range of Independence
Support		11. Level of Support

Figure B1: Ullman and D'Ambrosio's (1995) decision problem taxonomy.

The taxonomy lists altogether eleven categories. Eight of them are associated with the 'structure' of a decision situation. The aspects of 'structure' are 'decision space', 'preference model', and 'belief model'. The remaining three categories refer to 'focus', 'range', and 'support' of/for a decision situation. Ullman and D'Abrosio describe their categories as follows:

Category 1 - problem completeness

Information regarding the decision space may be complete or incomplete. Complete information means that all alternatives and all criteria are known at the time the evaluation is made. Incomplete information means that alternatives or criteria may evolve during further deliberations of the issue.

Category 2 - abstraction level

The alternatives and criteria may be described in a refined or in an abstract way. Refined information is quantitative; abstract information is qualitative.

Category 3 - determinism

The information regarding the alternatives and criteria may be deterministic, i.e. point-values, or distributed, i.e. value ranges.

Category 4 - objective function (effectiveness function)

An objective function is a mechanism for evaluating how well an alternative meets the entire set of criteria. This means, an objective function aggregates evaluation statements about the effectiveness of a particular alternative with respect to a single criterion into an overall evaluation statement for this alternative with respect to the entire set of criteria. In some cases the objective function will reflect the desire to minimise, maximise, or to find an optimum point for some parametric characteristics of the alternatives. In other cases, as in selection, the objective function is a weighted preference of criteria that call for judgement or saticficing.

Category 5 - consistency

If a decision is made by a group there may be different viewpoints on the importance of criteria and the evaluation of alternatives, i.e. preference models will vary. If there is only a single viewpoint, as in individual decision-making, the preference model is consistent. If there are many viewpoints the preference model is inconsistent.

Category 6 - comparison basis

The comparison basis is the type of comparison made in the evaluation of an alternative. An alternative may either be assessed absolutely on an independent scale with respect to a criterion or relatively by comparing the two alternatives with respect to a criterion.

Category 7 - dimension

The dimension of a decision situation refers to statements on the level of knowledge about an alternative with respect to a particular attribute and the level of confidence referring to the evaluation of an alternative with respect to a particular attribute. If the decision-makers neither states their level of knowledge nor their level of confidence the decision situation is zero-dimensional. If either their level of knowledge or their level of confidence are expressed the situation is one-dimensional, and if both, their level of knowledge and their level of confidence are expressed the situation is two-dimensional.

Category 8 - belief completeness

In a decision-making group every member can contribute a belief model, i.e. knowledge and confidence, towards the alternatives' evaluation. If all group members evaluate every alternative with respect to every criterion the group's belief model is complete. If this is not the case the group's belief model is incomplete.

Category 9 - problem focus

The problem focus depends on the issue being resolved. It may either be an issue on a product or on a process that supports the product.

Category 10 - range of independence

The range of a design issue's independence can be classified by three different types. A 'type 1' issue is completely independent on other issues; the relevant information is static. A 'type 2' issue is dependent on the resolution of other issues. Before a decision on a type 2 issue can be made its parent issues may have to be resolved. A 'type 3' issue is interdependent with other issues. A type 3 issue does not represent a linear relationship to one other issue, but it focuses on decomposition of issues into sub-issues and the subsequent composition of information derived in the sub-issues back into the higher level issue.

Category 11 - level of support

The level of support required for a decision-making problem may be classified as representation, outcome determination, or decision analysis. 'Level 1' support is limited to the representation of the issues, alternatives, arguments and criteria on which the decision is based. 'Level 2' support, outcome determination, helps the decision-maker to determine the effectiveness of a particular alternative with respect to a criterion. 'Level 3' support, decision analysis, determines which alternative to choose in order to satisfy an issue.

After studying design processes, Ullman and D'Ambrosio suggested that a typical decision situation in conceptual design may be characterised by their taxonomy as shown in figure B2. By using the taxonomy this way, it indicates requirements for design decision-making support. This means that a design decision-making method or tool should cope with a situation as described by the 'classification state' as in figure B2. For example, the method would be required to cope with an incomplete alternative space and with qualitatively rather than quantitatively described alternatives and criteria.

Alternative space	Completeness	Incomplete
	Abstraction level	Qualitative
	Determinism	Undefined
Preference model	Objective function	Judgement
	Consistency	Inconsistent
	Comparison basis	Absolute
Belief model	Dimension	Knowledge and confidence
	Completeness	Incomplete
	Focus	Product
	Range of independence	Interdependent
	Level of support	Decision analysis

Figure B2: Characteristics of decisions in conceptual engineering design.

Dwarakanath's study

Dwarakanath studied design processes and made a number of particular observations. These lead to a set of suggested requirements for decision-making support:

A) Support regarding the process of decision-making:

- Facilitating a systematic process by:
  - presenting sub-issues and interdependent issues,
  - showing weak and strong points of an alternative,
  - allowing for a flexible use of methods,
  - checking for consistency regarding the consideration of all criteria,
  - identifying unresolved issues,
  - presenting factors that dominate the decision,
  - structuring the decision-making process,
  - presenting all alternatives along with their respective arguments (strength, weaknesses).
- Raising awareness by:
  - presenting a quick overview over the process stages,
  - presenting relationships between decisions,
  - presenting previous decisions along with relevant information.

B) Support regarding the generation of documentation

- Foster generation of structured documentation about the decision-making process and design history by:
  - capturing design issues, alternative solutions, arguments generated, criteria established and decisions made,
  - organising information in a structured way to keep track of the relationships between different types of captured information.

C) Support regarding the reuse of information:

- Providing a convenient way of retrieving captured information.

D) Support regarding communication within teams:

- Providing information simultaneously to all designers involved in a team. In particular:

- keeping track of decisions made by different team members,
- communicating newly captured information automatically to all team members.

E) Support regarding the use of methods and tools:

- Suggesting suitable methods and tools for the evaluation of alternatives;
- Creating input to chosen methods and tools automatically.

Based on the requirements identified through his study, Dwarakanath proposed a decision support framework. This framework is shown in figure B3. The core of the framework consists of the three stages: capture information, organise information, and present information. Basically, information is captured from the design process, organised appropriately, processed through evaluation methods and eventually the results are presented back to designers who then progress with their design process.

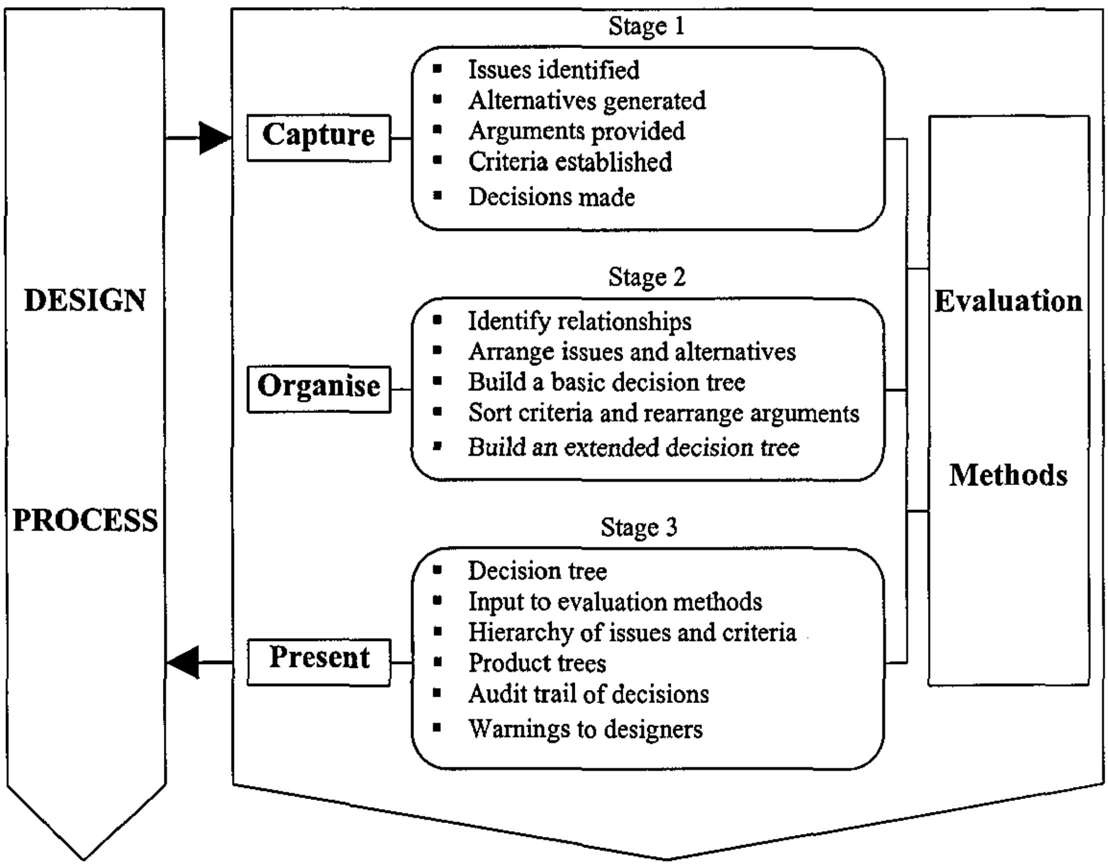


Figure B3: Dwarakanath's (1996) decision support framework.

Appendix C

Workshop Handout

Introduction to the design issue

Carbon Dioxide ($C0_2$) is used to pressurise barrels of beer. The $C0_2$ gives the beer a head and, in some cases, helps to force it from the barrel when it is served. The $C0_2$ is stored in steel cylinders that are delivered to the Public House (pub) by lorry. The cylinders are then connected to the barrel by means of a system of pipes end valves.

Over 60% of pubs in the United Kingdom store their beer barrels and gas cylinders in a cellar below ground level. Access to the cellar is usually by means of internal stairs and a trap door located outside the external walls of the pub. The cylinders are delivered to the pub on a lorry. The driver/operator lowers the required number of cylinders to ground level using an electric tailgate on the lorry. If the lorry cannot park near to the trap door, the cylinders are then loaded on a ‘sack truck’, which is also stored on the lorry. The cylinders are then taken to the trap door entrance to the cellar. With the trap door raised, the operator will have a view into the cellar and of the barrel role that is present in most cases. The operator now needs to safely lower the cylinders into the cellar. ‘Safely’ refers to the operator, the staff and customers of the pub and members of the general public. This latter requirement is particularly relevant because the trap door may be located in the car park of the pub or on the pavement of a public road. All relevant codes of practice and health and safety requirements must be adhered to.

A device is required that enables the operator to place the full gas cylinders into the cellar and raise the empty cylinders. The device should be portable; it is not intended to be in permanent position at the pub.

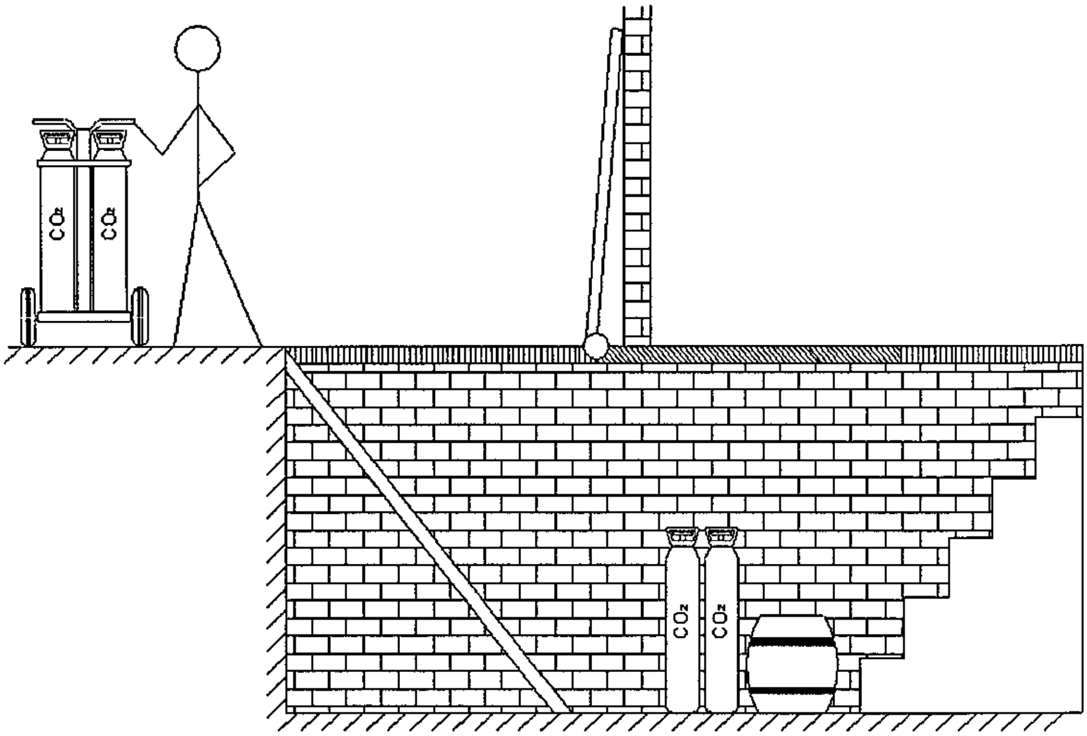
Product Design Specification

Item	Description	Details
1	Cylinder mass	16kg (empty) 22kg (full)
2	Cylinder height	940mm
3	Cylinder diameter	140mm
4	Cylinder fittings	Valve enclosed in plastic protective collar.
5	Cellar height (floor to ceiling)	2m (1.8m to 2.5m range)
6	Angle of barrel role	45 degrees (typical)
7	Clear floor area at base of barrel role	2 m^2
8	Trap door dimensions	1.5m x 2.00m (typical
9	Trap door attachment arrangements	Hinge at back or lift off

The number of cylinders delivered at a time to each pub varies between 2 and 18, although the average is 6. Currently, the delivery-time for 6 cylinders is approximately 35 minutes. This would comprise 10 minutes talking to the pub owner and obtaining signatures, 10 minutes taking the cylinders from the lorry to the cellar door, 10 minutes taking them into the cellar and 5 minutes stowing the sack truck on the lorry and re-securing the load.

The driver of the lorry is not supposed to ask the publican for assistance because of legal consequences in case of an accident. However, this rule is frequently broken. In addition to the cylinders described above, the driver also delivers larger (32kg, 870mm high, 203mm diameter) cylinders of nitrogen in the same way. These present an even greater risk of injury, but are not to be included in this study. The device should reduce the risk of injury to the driver by removing the need to lift and sustain heavy loads in 'difficult' positions. The device must be commercially viable in terms of initial cost, running cost and its effect on the efficiency of the delivery service.

Diagrammatic general layout of pub and cellar



Criteria for evaluating the alternative design solutions

Device for lowering and lifting beer gas cylinders

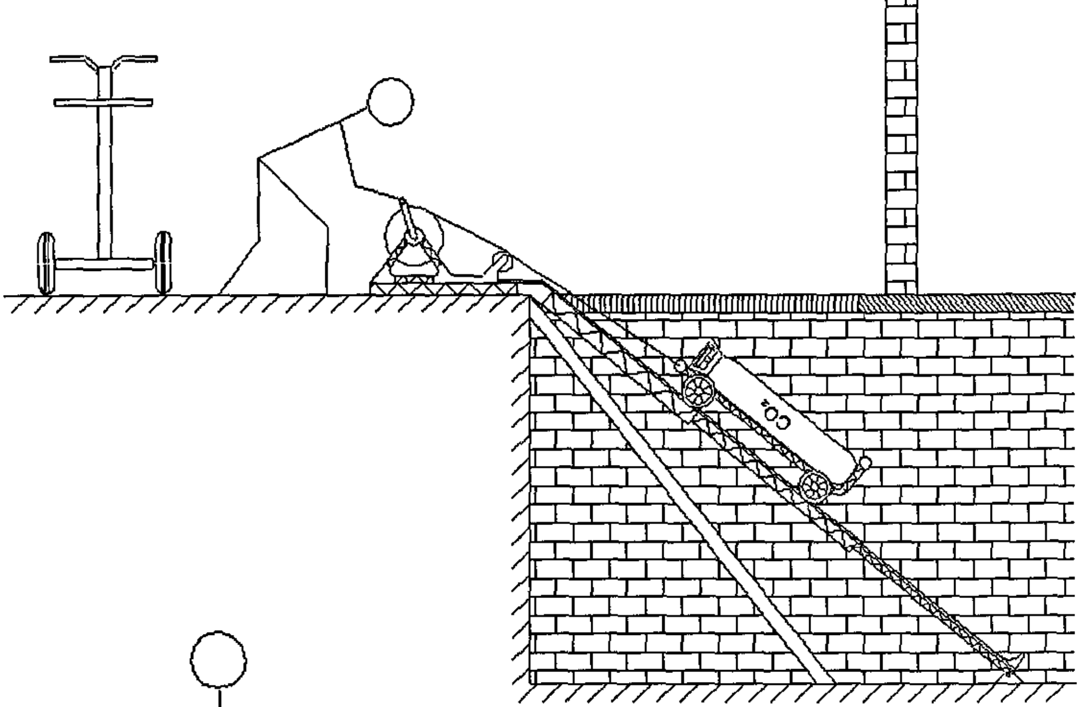
- Safety
  - For operator
  - For staff of pub
  - For customers of pub
  - For general public
- Efficiency of delivery
- Cost
  - Initial cost
  - Running cost
  - Maintenance cost
    - Maintenance intervals
    - Requirement of special tools
- Ease of use
  - Force required for operation
  - Training required for operation
  - Weight to be handled
- Versatility
  - Different Cylinders
  - Different cellars
    - Drop depth
    - No barrel role
    - Size of trap door

Alternative design solutions (concepts)

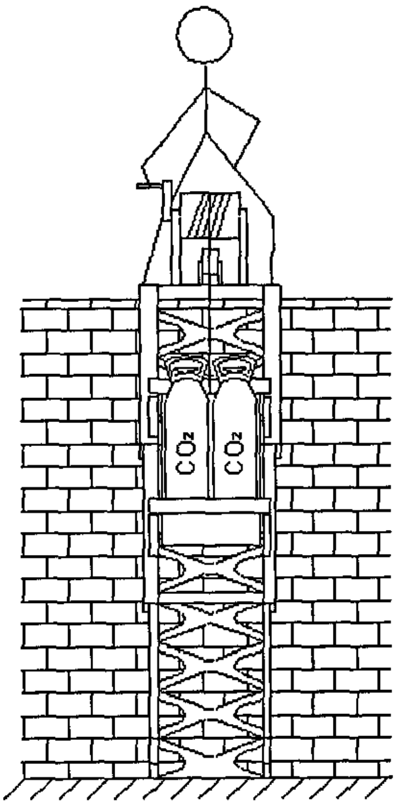
A: Telescope-rail-bound carriage

A telescope ladder creates a pair of rails on which a carriage holding the gas cylinders can be moved upwards and downwards. The carriage's movement may be controlled by a rope or some mechanism.

Side view

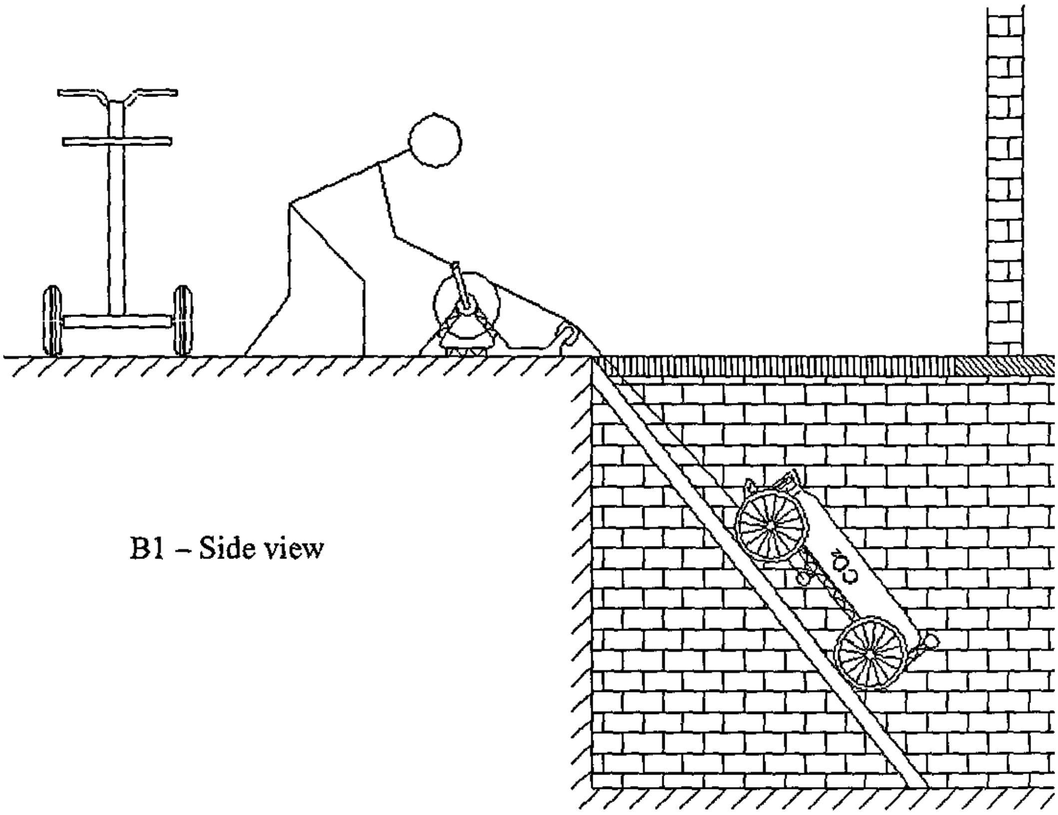


Front view

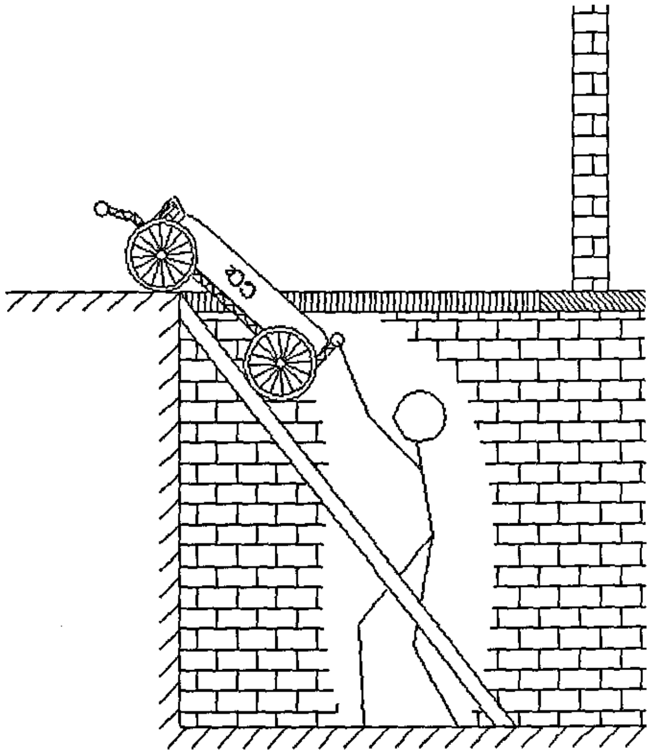


B: Carrier guided by barrel role

This alternative is similar to concept A. However, no telescope ladder is needed. Instead, the barrel role, if available, is used as guide rails for the carriage. In B1 the carriage is lead by a rope. In B2 the carriage's speed is controlled by an automatic brake system in the wheels. In B2 the carriage must be pushed upwards manually.



B1 – Side view

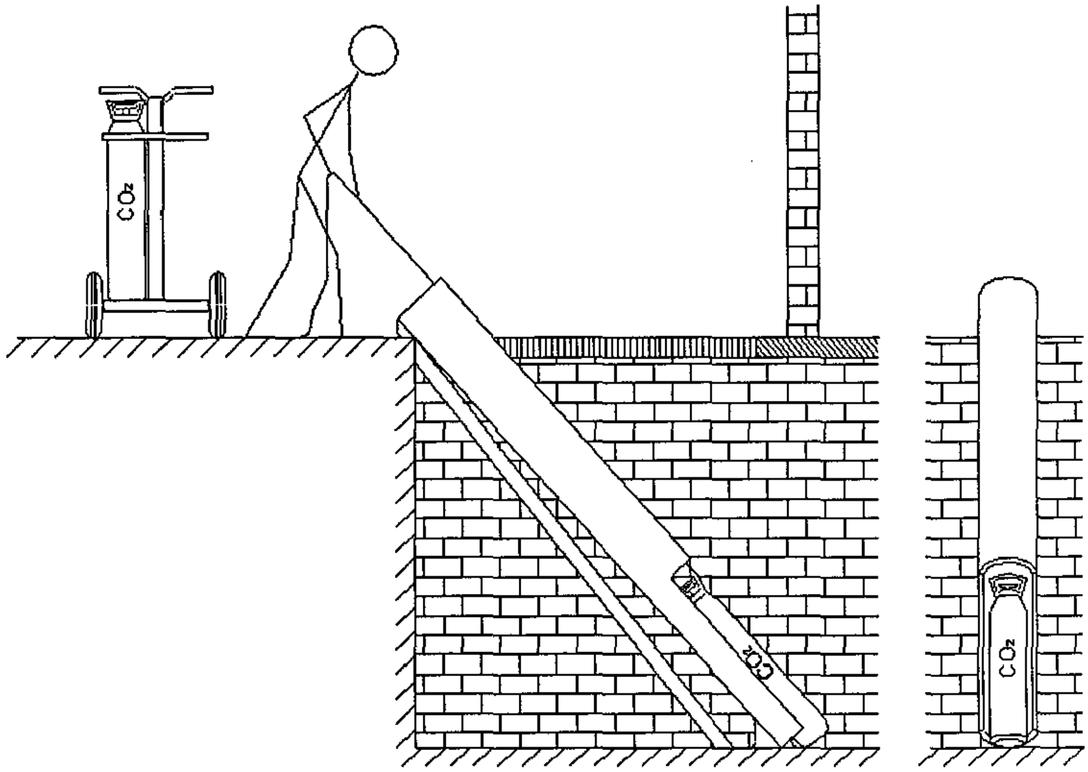


B2 – Side view

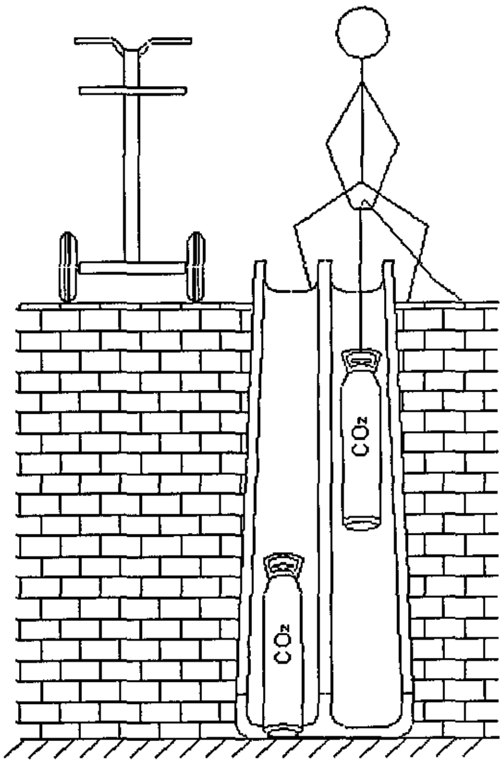
C: Tube or slide

The gas cylinders move freely down a tube or slide. They can be pulled upwards by a rope.

C1 - Tube, side and front view:

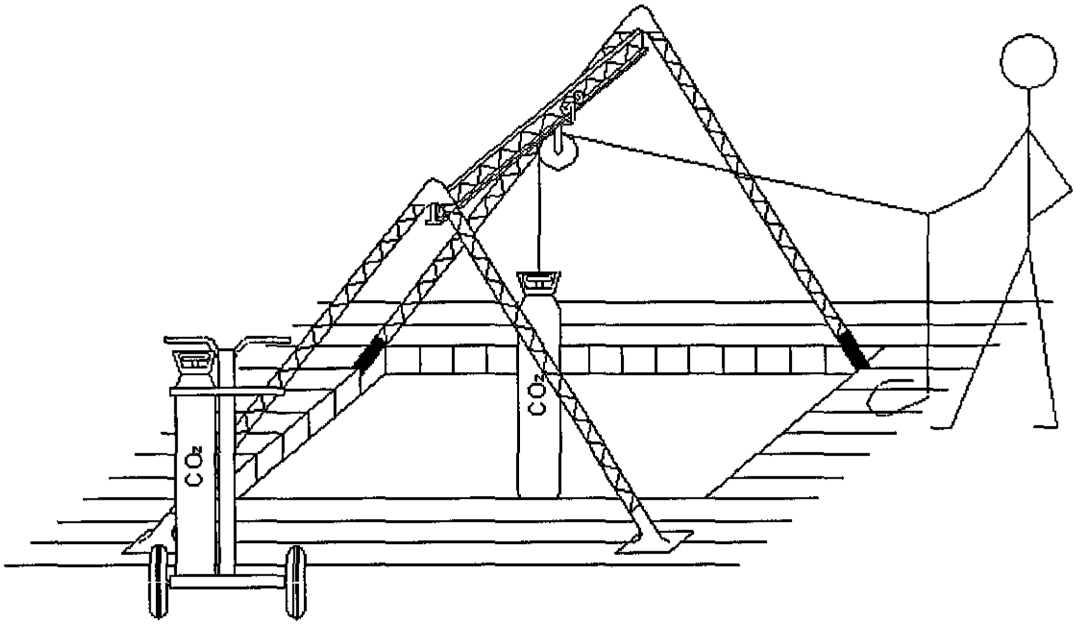


C2 - Slide, front view:



D: Pulley at framework

The gas cylinders are attached to a rope, which is guided by a pulley. The pulley can be moved horizontally at a framework above the trap door.



Appendix D

Observed Decision-Making Processes

Within the main text the three groups' decision-making processes were shown graphically as flow charts (see figures 4.1, 4.2 and 4.3). Within this appendix the same processes are outlined in more detail in form of tables. These tables contain four columns: step, counter/timer, aspect addressed and remarks.

The column 'step' indicates process steps in accordance with my decision-making process model as developed in section 4.4 and shown graphically in figure 4.22. Please note that not all groups went through all process steps.

The column 'counter/timer' indicates counter or timer positions. For generating the real-time protocols I used a video camera and two different types of audio recorders. As the audio quality of the video recordings did hardly allow for producing exact transcripts (see also section 3.5.2), I generated them almost entirely by using the audio recorders. For recording and transcribing the two student groups I used tape recorders that had in-built *counters*, but no timers. For recording and transcribing the professional group I used a mini-disc recorder that had an in-built *timer*. Hence, the two tables on the students' decision-making processes show 'counter positions', whereas the table on the professionals' decision-making process shows 'timer positions'.

The column 'aspect addressed' indicates various aspects of decision-making processes as addressed by the groups (see section 3.4.6).

The column 'remarks' contains some comments on the processes.

The table *rows* are on 'meta-construct level' (see section 3.4.6). This means that various sections (see figure 3.4) are summarised to larger constructs for a better overview and concise presentation.

Decision-making process of the informal students

Step	Counter	Aspect addressed	Remarks
1	0-85	Determine general evaluation approach	They agreed to evaluate each alternative with respect to each criterion in turn (criterion based evaluation)
2B.2	85-218	Evaluate alternative restrictedly; Concept A with respect to criterion: ▪ safety	Criteria-based approach
2A	218-323	Evaluate alternative comprehensively; Concept A; Advantages with respect to: ▪ ease of use ▪ versatility - different cellars Disadvantage with respect to: ▪ versatility - different cellars	Switch to comprehensive approach Criteria are mostly not explicitly mentioned; All criteria are treated as if they were equally important.
2B.2	323-664	Evaluate alternative restrictedly; Concept A with respect to criteria: ▪ safety ▪ cost ▪ running cost ▪ maintenance cost ▪ requirement of special tools ▪ efficiency - time for set-up ▪ safety ▪ versatility - different cellars ▪ efficiency of delivery	Criteria-based approach is taken up again Criterion 'safety' is considered second time Criterion 'safety' is considered third time
1	664-805	Discussion on approach for further evaluations	It was suggested to use the comprehensive approach for the remaining alternatives
2A	805-890	Evaluate alternative comprehensively; Concept A; Advantages with respect to: ▪ ease of use ▪ efficiency of delivery	Comprehensive approach
2A	890-973	Evaluate alternative comprehensively; Concepts B1 and B2; Disadvantages with respect to: ▪ versatility ▪ cost ▪ efficiency of delivery - time for set up	Comprehensive approach
2A	973-997	Evaluate alternative comprehensively; Concept C1; Disadvantages with respect to: ▪ ease of use ▪ safety	Comprehensive approach

Step	Counter	Aspect addressed	Remarks
2A	997-1070	Evaluate alternative comprehensively; Concept C2; Disadvantages with respect to: ▪ ease of use ▪ safety Advantage with respect to: ▪ efficiency of delivery	Comprehensive approach
2A	1070-1181	Evaluate alternative comprehensively; Concept D; Disadvantage with respect to: ▪ efficiency of delivery	Comprehensive approach
3.1	1181-1236	Establish ranking order	Only rank 1 and 2 are justified relatively to each other
3.1	1236-1503	Analyse ranking	They deliberated various sub-issues regarding the highest ranked alternative, i.e. concept A
3	1503-1528	Concluding remarks and termination of concept selection process	

Decision-making process of the formal students

Step	Counter	Aspect addressed	Remarks
1	0-46	Determine general evaluation approach	
2B.1	46-158	Structure criteria	This structure looked almost exactly as the criteria tree given to the group
2B.1	158-356	Differentiate criteria into constraints and objectives	Some criteria are re-formulated to make implicit constraints explicit. The constraints were derived exclusively from the sub-criteria under the label 'safety' and in particular from 'safety – for operator' The first objective was derived from the given criteria 'safety – for staff', 'safety – for customers', and 'safety – for general public'. The remaining objectives associate to the given criteria
1	356-470	Refine the evaluation approach	It was suggested to use an evaluation matrix, which was then drawn up, a scoring scale was chosen and a scale for criteria-importance weightings was agreed upon

Step	Counter	Aspect addressed	Remarks
2B.1	470-558	Establish criteria-importance weightings for objectives: ▪ hazard to staff, customers, … ▪ initial cost ▪ running cost ▪ efficiency of delivery ▪ initial cost ▪ efficiency of delivery ▪ versatility ▪ different cellars ▪ different cylinders ▪ ease of use - force required ▪ ease of use - training required	Some goals are considered more than one time. Previously established importance weightings are revisited and changed. Thus, assigning these weightings is an iterative process. The weightings are relative. In some cases the weights were intuitively assigned without any rational justification but in other cases the weights caused extensive discussions on importance and meaning of the objective in question. In these cases the rationales for particular weights were expressed explicitly.
2B.2	558-587.5	Check alternative towards constraint satisfaction Concept A: ▪ all constraints	
2B.2	587.5-604.5	Check alternative towards constraint satisfaction Concept B1: ▪ all constraints	
2B.2	604.5-620	Check alternative towards constraint satisfaction Concept B2: ▪ all constraints	
2B.2	620-659.5	Check alternative towards constraint satisfaction Concepts C1 and C2: ▪ all constraints	Concepts C1 and C2 are addressed simultaneously
2B.2	659.5-677	Check alternative towards constraint satisfaction Concept D: ▪ all constraints	
1	677-716	Further refine evaluation approach	They discussed the details of how to implement a pair-wise comparison method and they chose to use concept A as the datum.
2B.2	716-821	Evaluate alternatives restrictedly with respect to criterion 'hazards to staff, customers or general public within the vicinity'; Concepts: ▪ B1 ▪ B2 ▪ C1 ▪ C2 ▪ D	

Step	Counter	Aspect addressed	Remarks
2B.2	821-846	Evaluate alternatives restrictedly with respect to criterion 'initial cost'; Concepts: ▪ B1 and B2 ▪ C1 ▪ C2 ▪ D	Concepts B1 and B2 are now considered together.
2B.2	846-871	Evaluate alternatives restrictedly with respect to criterion 'running cost'; Concepts: ▪ B1 and B2 ▪ C1 and C2 ▪ D	Concepts C1 and C2 are now also considered together.
2B.2	871-890.5	Evaluate alternatives restrictedly with respect to criterion '*maintenance cost*'; Concepts: ▪ B1 and B2 ▪ C1 and C2 ▪ D	
2B.2	890.5-905	Evaluate alternatives restrictedly with respect to criterion 'efficiency of delivery'; Concepts: ▪ B1 and B2 ▪ C1 and C2 ▪ D	
2B.2	905-920	Evaluate alternatives restrictedly with respect to criterion 'versatility - different cylinders'; Concepts: ▪ B1 and B2 ▪ C1 ▪ C2 ▪ D	Concepts C1 and C2 are considered separate again
2B.2	920-924	Evaluate alternatives restrictedly with respect to criterion 'versatility - different cellars'; Concepts: ▪ B1 and B2 ▪ C1 ▪ C2 ▪ D	
2B.2	924-982	Evaluate alternatives restrictedly with respect to criterion 'ease of use - force required'; Concepts: ▪ B1 ▪ B2 ▪ C1 ▪ C2 ▪ D	Concepts B1 and B2 are also considered separate again. It is found that concept B2 does not actually satisfy the constraint 'no awkward lifting of anything over 15kg'. Yet, it was decided to see how this concept performs comprehensively. In case its comprehensive performance is extremely good it will be changed in such a way that it is going to meet the constraint.

Step	Counter	Aspect addressed	Remarks
2B.2	982-993	Evaluate alternatives restrictedly with respect to criterion 'ease of use - training required'; Concepts: ▪ B1 ▪ B2 ▪ C1 and C2 ▪ B2 ▪ D	Alternative B2 is evaluated twice.
3.1	993-1044	Determine comprehensive performances for all alternatives and establish ranking	
3.2	1044-1081.5	Discuss comprehensive evaluation of, in particular, concepts: ▪ B2 ▪ C2 ▪ D ▪ A in comparison to C2 ▪ C1 in comparison to C2	
3	1081.5	Agree on final selection	

Decision-making process of the professionals

Step	Timer (min.sec)	Aspect addressed	Remarks
1	00.00-00.04	Determine general evaluation approach	
2B.1	00.04-09.40	Structure criteria	New criteria were identified and included in the tree structure.
1	09.40-18.00	Refine general evaluation approach	Agree on two-stage approach; define details about first stage: use of evaluation matrix and datum method, agree on scoring scale
2B.2	18.00-31.51	Evaluate alternative restrictedly; Concept A: ▪ all criteria	At a number of occasions it is considered to introduce a more diversified scoring scale
2B.2	31.51-37.37	Evaluate alternative restrictedly; Concept B: ▪ all criteria Concept A with respect to criteria: ▪ initial cost ▪ running cost ▪ weight to be handled	Due to the identification of relatively large performance differences the scoring scale was refined and retrospectively applied to some restricted evaluations of concept A. So, the comparison basis is now not only the datum, but also concept A.
2B.2	37.37-41.49	Evaluate alternative restrictedly; Concept C with respect to criterion: ▪ safety for operator	They decided that it would be more appropriate to evaluate concept C1 and concept C2 separately
2B.2	41.49-47.57	Evaluate alternative restrictedly; Concept C1:	Concepts A and B were re-evaluated with respect to the criterion 'running

Step	Timer (min.sec)	Aspect addressed	Remarks
		▪ all criteria Concepts A and B with respect to criterion: ▪ running cost	cost'. This was after concept C1 was considered with respect to this criterion
2B.2	47.57-52.34	Evaluate alternative restrictedly; Concept C2: ▪ all criteria Concept C1 with respect to criterion: ▪ time of each delivery	Concept C1 was re-evaluated with respect to criterion 'time of each delivery'
2B.2	52.34-57.31	Evaluate alternative restrictedly; Concept D: ▪ all criteria	
1	57.31-51.17	Reflect on evaluation approach	The new criterion 'ease of product development' was identified. Yet, it was ignored in the further evaluation process
3.1	61.17-63.20	Determine comprehensive performances for all alternatives and establish ranking	
3.2	63.20-69.10	Analyse ranking with respect to all concepts, but with particular emphasis on: ▪ concept C2 ▪ concepts D compared to E ▪ concepts B compared to A ▪ concepts B compared to C1 ▪ concepts B, C1, compared to C2 ▪ concepts A, B, C1, compared to C2 ▪ concepts B, C1, compared to C2	The group deliberated the comprehensive ranking. They tried to understand the particular reasons for the ranking positions of specific alternatives. Finally, concepts B and C2 were pre-selected
1	69.10-72.10	Refine general approach for further evaluations	
2B.2	72.10-81.59	Evaluate alternatives restrictedly; Concepts B and C2: ▪ all criteria	The two pre-selected alternatives were evaluated together with respect to each criterion in turn. The new criterion 'space requirement on lorry' was identified. Yet, it was ignored in the further evaluation process
3.1	81.59-82.38	Determine comprehensive performances for concepts B and C2 and establish ranking	
3.2	82.38-82.53	Analysis of calculated ranking and termination of concept selection process	

Appendix E

Transcribed Activity Patterns

Within the main text three models of specific activity patterns were introduced. These were (i) a pattern related to process step 2A.2 'support comprehensive evaluation' (see figure 4.13), (ii) a pattern related to process step 2B.1.3 'determine importance of criterion (see figure 4.16) and (iii) a pattern related to process step 2B.2A/B.3 'evaluate alternative-criterion pair' (see figure 4.18). This appendix shows various excerpts from the transcripts as examples for the mentioned activity patterns.

Support comprehensive evaluation

Below are two transcript excerpts showing activity patterns related to the process steps 2A.1 and 2A.2. They are both taken from the transcript of the informal students' decision-making process. The different shades indicate that the two examples represent different constructs (of the same type, see section 3.4.6 and figure 3.4)

First excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
218	B: What do you think of the idea of concept A as a whole?	Controls the process by prompting for an evaluation of concept A	Evaluation of concept A	Controlling the process
222	A: I think it's pretty alright ...	Evaluates concept A comprehensively	Evaluation of concept A; Evaluate alternative comprehensively	Evaluating comprehensive performances informally
225	A: ... because one thing ... it's transportable ... okay.	Assesses concept A's performance restricted to criterion 'ease of use'	Evaluation of concept A; Support comprehensive evaluation	Determining or evaluating restricted performances informally
231	C: ...it's compact, isn't?	Delivers some evidence for 'transportability' and hence, 'ease of use'	Evaluation of concept A; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
232	A: Yeah, it's compact.	Agrees with above evidence for 'transportability' and hence, 'ease of use'	Evaluation of concept A; Support comprehensive evaluation	Raising evidence on restricted performance of concepts

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
238	A: ... it looks a simple design ...	Raises evidence for the assessment of concept A with implicit respect to the criterion 'ease of use' or 'cost'	Evaluation of concept A; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
245	A: I mean, it really depends on how they design this ... whether it is geared, whether there is any mechanism to lower this (the carrier) or is it just ... to slight down ...	Discusses the sub-issue 'how to operate the carrier'	Evaluation of concept A; Support comprehensive evaluation	Discussing sub-issues
251	C: The thing is ... obviously, you are going to have different angles, aren't you? You can have it at different angles so that the width ... to the operator restricted from falling down is going to be related to the angle, that's it, isn't it? So if the basement of the cellar is really small, ... from the door ... there isn't that much room.	Raises evidence for the performance of alternative A with respect to criterion 'different cellars'	Evaluation of concept A; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
275	C: You just gonna have all the problems with actually lowering it down at ... some cellars.	Evaluates the performance of concept A with respect to criterion 'different cellars'	Evaluation of concept A; Support comprehensive evaluation	Determining or evaluating restricted performances informally
288	B: It's quite good ...	Evaluates concept A comprehensively	Evaluation of concept A; Evaluate alternative comprehensively	Evaluating comprehensive performances informally
294	B: ... for it stands alone as well.	Raises evidence for following performance assessment of concept A with respect to criterion 'different cellars' (explains why a barrel role is not necessary)	Evaluation of concept A; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
296	B: You don't have to have a barrel role to use it - that's good.	Evaluates the performance of concept A with respect to criterion 'different cellars'	Evaluation of concept A; Support comprehensive evaluation	Determining or evaluating restricted performances informally
298	B: But then again, why wouldn't they have a barrel role into a cellar. How would they get the barrels down otherwise?	Discusses environment, i.e. current cellar structures	Evaluation of concept A; Support comprehensive evaluation	Determining the product environment
304	A: There might be some (cellars) ... you (the researcher) did mention there might be some without a barrel role, right?	Discusses environment, i.e. current cellar structures	Evaluation of concept A; Support comprehensive evaluation	Determining the product environment
308	Researcher: There are some, yes.	Discusses environment, i.e. current cellar structures	Evaluation of concept A; Support comprehensive evaluation	Gaining external information

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
316	B: Maybe there are mats at the bottom of the cellar.	Discusses environment, i.e. current cellar structures	Evaluation of concept A; Support comprehensive evaluation	Determining the product environment
318	C: It doesn't matter.	Controls evaluation process by stopping an irrelevant discussion	Evaluation of concept A;	Controlling the process
323	BREAK			
382	C: ... but then, is it safe? If it (the telescopic rails) just slights down, is it going to be safe for any staff of the pub down in the cellar?	Explains (raises evidence) why there might be a safety problem with concept A	Evaluation of concept A; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
390	A: Let's assume this is a geared design, because it (the carrier) is not able to go up. So, there will be some mechanism to pull up this (the carrier) ...	Makes an assumption on a solution for sub-issue 'how to operate carrier' with respect to concept A	Evaluation of concept A; Support comprehensive evaluation	Accepting assumptions on sub-issue solutions
400	B: Why? It could be done with a bit of rope. As long as there is a rope attached to it here (at the carrier) ... all you need is a person standing there (at the top) to raise and lower it.	Deliberates sub-issue 'how to operate carrier' with respect to concept A	Evaluation of concept A; Support comprehensive evaluation	Discussing sub-issues
410	A: But then it would be the same as this design (the carrier on the barrel role). I don't think this is the idea of ...	Deliberates sub-issue 'how to operate carrier' with respect to concept A	Evaluation of concept A; Support comprehensive evaluation	Discussing sub-issues
417	B: The difference is, it's got rails, it doesn't have to rely on the barrel role. It says (in the design description), it is similar to A (the telescope design).	Explains the working principle of concept A and concept B;	Evaluation of concept A; Support comprehensive evaluation	Clarifying concept working principle
426	B: But yeah, you can just lower it up and down on a rope. You don't have to have mechanisms. That's just going to boost the cost up. It's only something to get the cylinders up and down.	Deliberates sub-issue 'how to operate carrier' with respect to concept A	Evaluation of concept A; Support comprehensive evaluation	Discussing sub-issues
435	A: So okay, they gonna use a rope ...	Makes an assumption on a solution for sub-issue 'how to operate the carrier' with respect to concept A	Evaluation of concept A; Support comprehensive evaluation	Accepting assumptions on sub-issue solutions
439	A: ... the safety part ... there might be a problem on the safety part, right?	Assesses the performance of concept A with respect to the criterion 'safety'	Evaluation of concept A; Support comprehensive evaluation	Determining or evaluating restricted performances informally

Second excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
997	B: Slide!	Controls evaluation process by suggesting the consideration of another alternative	Evaluation of concept C2	Controlling the process
999	B: Okay ... well ... no ... (gestures disapproval)	Indicates that overall the alternative is not seen satisfactory	Evaluation of concept C2; Evaluate alternative comprehensively	Evaluating comprehensive performances informally
1002	B: ... it can do more than one (cylinder at a time) ...	Raises evidence for an assessment of concept C2 with respect to the criterion 'efficiency of delivery"	Evaluation of concept C2; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
1004	B: ... but then they may bang against each other, bang against the side, bang against the bottom ...	Raises evidence for an assessment of concept C2 with respect to the criterion 'safety'	Evaluation of concept C2; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
1007	C: And when they get to the bottom they may roll in the back of the cellar as well.	Raises evidence for an assessment of concept C2 with respect to the criterion 'safety'	Evaluation of concept C2; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
1014	C: You need some kind of restraint on the bottom there to stop it get any further.	Deliberates solution for sub-issue 'how to stop cylinders at bottom'	Evaluation of concept C2; Support comprehensive evaluation	Discussing sub-issues
1020	C: Yeah, I would say its biggest downfall is its size ...	Assesses the performance of concept C2 with respect to criterion 'ease of use'	Evaluation of concept C2; Support comprehensive evaluation	Determining or evaluating restricted performances informally
1023	C: ... and the fact that it's not very safe.	Assesses the performance of concept C2 with respect to criterion 'safety'	Evaluation of concept C2; Support comprehensive evaluation	Determining or evaluating restricted performances informally
1026	C: It's efficient ...	Assesses the performance of concept C2 with respect to criterion 'efficiency of delivery'	Evaluation of concept C2; Support comprehensive evaluation	Determining or evaluating restricted performances informally
1029	C: ... provided you could get to install this quick enough. You will have to get it right the way off the lorry and get it down, but after that ...	Raises evidence for an assessment of concept C2 with respect to the criterion 'efficiency of delivery'	Evaluation of concept C2; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
1032	B: I was wondering what this is made of ...	Deliberates sub-issue 'material' for alternative C2 in association to earlier considered aspect weight referring to 'ease of use'	Evaluation of concept C2; Support comprehensive evaluation	Discussing sub-issues
1038	B: ... you would probably have this made of plastic.	Deliberates sub-issue 'material' for concept C2 and assumes plastic as material	Evaluation of concept C2; Support comprehensive evaluation	Accepting assumptions on sub-issue solutions
1040	A: Yeah.	Deliberates sub-issue 'material' for concept C2 and supports above assumption	Evaluation of concept C2; Support comprehensive evaluation	Accepting assumptions on sub-issue solutions

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
1042	A: Yeah, it is light, but the thing is, it is bulky, it's huge ...	Raises evidence for assessment of concept C2 with respect to the criterion 'ease of use'	Evaluation of concept C2; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
1044	A: ...so, it's quite difficult to handle.	Assesses the performance of concept C2 with respect to criterion 'ease of use'	Evaluation of concept C2; Support comprehensive evaluation	Determining or evaluating restricted performances informally
1046	C: Yeah.	Agrees with assessment above	Evaluation of concept C2; Support comprehensive evaluation	Determining or evaluating restricted performances informally
1048	A: It's no matter of weight, it's more of size.	Introduces a new aspect on the criterion 'ease of use': 'size to be handled' in addition to 'weight to be handled'	Evaluation of concept C2; Support comprehensive evaluation	Identifying criteria
1050	B: If you've got a lorry backed right up, you just tip it out, couldn't you? Pull the end out, tip it down and it's in there (indicates: this is a question to the researcher on whether the environmental conditions would allow for this)	Attempts to raise evidence for an assessment of concept C2 with respect to criterion 'ease of use' (see 1026)	Evaluation of concept C2; Support comprehensive evaluation	Raising evidence on restricted performance of concepts
1055	Researcher: (mentions that many trap doors are, according to the description, not directly accessible by road)	Provides information on product environment	Evaluation of concept C2; Support comprehensive evaluation	Gaining external information
1063	B: I'd say we don't like that one then.	Comprehensively evaluates concept C2	Evaluation of concept C2; Support comprehensive evaluation	Evaluating comprehensive performances informally

Determine importance of criterion

Below are two transcript excerpts showing activity patterns related to the process step 2B.1.3. They are both taken from the transcript of the formal students' decision-making process. The different shades indicate that the examples represent different constructs (of the same type, see section 3.4.6 and figure 3.4).

First excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
485.5	C: 'Different cylinders', ...	Controls evaluation process by suggesting the consideration of another criterion;	Determine importance of criterion; Criterion 'versatility - different cylinders'	Controlling the process
486	C: ... 'different cylinders' is five.	Suggests an importance weighting for the criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Weighting criteria formally
486.5	A: How many different cylinders are there?	Tries to clarify the product environment	Determine importance of criterion; Criterion 'versatility - different cylinders'	Clarifying the product environment
487.5	E: You may end up just having different attachments for different sizes.	Raises evidence for the below criteria weighting for the criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Raising evidence on criteria weights
488	E: So ... I think 'different cellars' is more important than 'different cylinders'.	Deliberates the importance of the criteria 'different cellars' and 'different cylinders'.	Determine importance of criterion; Criteria 'versatility - different cellars' and 'versatility – different cylinders'	Weighting criteria informally
488.5	B: Yeah, I think so too.	Agrees with above deliberations on the importance of criteria 'different cellars' and 'different cylinders'	Determine importance of criterion; Criteria 'versatility - different cellars' and 'versatility – different cylinders'	Weighting criteria informally
489	C: Alright, so make it five for 'different cellars'.	Suggests an importance weighting for the criterion 'different cellars'	Determine importance of criterion; Criterion 'versatility - different cellars'	Weighting criteria formally
489.5	E: Yeah.	Agrees with above suggestion on the importance weighting of criterion different cellars'	Determine importance of criterion; Criterion 'versatility - different cellars'	Weighting criteria formally
490	C: The thing is ... we don't have any ... we are not told anything about different cylinders, are we?	Deliberates the product environment regarding possible cylinder sizes	Determine importance of criterion; Criterion 'versatility - different cylinders'	Clarifying the product environment
492.5	E: Hmm, if it's a tube, it can only take one size, can't it?	Raises evidence for the below criteria weighting of the criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Raising evidence on criteria weights

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
495	A: So, 'different cylinders' ... rate?	Controls the evaluation process by urging the group to generate a weighting	Determine importance of criterion; Criterion 'versatility - different cylinders'	Controlling the process
495.5	C: I think 'different cylinders' is quite important.	Suggests an informal importance weighting for the criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Weighting criteria informally
496	E: Yeah, make it five.	Agrees with the above suggestion and specifies a formal weighting for the criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Weighting criteria formally
496.5	D: No, not five.	Disagrees with above suggestion on importance weighting for criterion 'different cylinders' and suggests another importance weighting (see 488)	Determine importance of criterion; Criterion 'versatility - different cylinders'	Weighting criteria formally
497	E: Okay, make it four.	Agrees with above suggestion on importance weighting for criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Weighting criteria formally
497.5	A, C, D: yeah.	Agree with above suggestion on importance weighting for criterion 'different cylinders'	Determine importance of criterion; Criterion 'versatility - different cylinders'	Weighting criteria formally

Second excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
498	E: 'Force required'?	Controls evaluation process by suggesting the consideration of another criterion	Determine importance of criterion; Criterion 'force required'	Controlling the process
498.5	A: Does that mean, as in lifting something?	Deliberates meaning of criteria 'force required'	Determine importance of criterion; Criterion 'force required'	Defining criteria
499	C: I think that's the whole point.	Raises evidence for the significance of criterion 'force required' (reminds the group on purpose of device)	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
499.5	E: That's the key think –	Raises evidence for the significance of criterion 'force required' (reminds the group on purpose of device)	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
500	E: ... you just can't be pulling more than ...	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
500.5	C: ... twenty kilogram ...	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
501	E: ... or fifteen if it's an awkward position.	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
502	D: We can constrain it, can't we? Just give it a number.	Deliberates meaning of criterion 'force required' and suggests it may be a constraint	Determine importance of criterion; Criterion 'force required'	Defining criteria
503	E: Hmm ... we can compare the forces involved ... because we have got different ways ...	Deliberates the practicality of criterion 'force required' and suggest a way for dealing with it	Determine importance of criterion; Criterion 'force required'	Defining criteria
511	C: So, what do you think, 'force required'?	Controls the evaluation process by stopping the discussion on the meaning of criterion 'force required' and by prompting the group members to give a statement on this criterion's importance weighting	Determine importance of criterion; Criterion 'force required'	Controlling the process
512	E: Fairly important.	Suggests an informal importance weighting for the criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria informally
512.5	E: If you are doing it all day then you don't want to be pulling lots of weight.	Raises evidence for the importance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
513	E: So, I would say at least four.	Suggests a formal importance weighting for the criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally
513.5	D: I would say, at least four or even five.	Suggests an importance weighting for the criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally
514.5	A: Hmm, go for four.	Partly disagrees with above statement on the importance of criterion 'force required' and suggests another importance weighting for this criterion	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally
515	C: The whole point is, one man only ...	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
516	D: Like he just said, if you are doing it all day ...	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
517	B: The whole point really is to meet the EU regulations, alright?	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
520	E: How can we assess the force and compare?	Deliberates meaning of criterion 'force required', in particular the means for its assessment	Determine importance of criterion; Criterion 'force required'	Defining criteria
521	C: Well, we multiply 20kg by 'g' and then we have a force, haven't we?	Deliberates meaning of criterion 'force required', in particular the means for its assessment	Determine importance of criterion; Criterion 'force required'	Defining criteria

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
526.5	E: We are talking about repetitive strain as well.	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
527	A: Yeah.	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
527.5	E: I mean, it is important ...	Suggests an informal weighting for the criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria informally
528	E: ... because this is the whole point of the thing.	Raises evidence for the significance of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Raising evidence on criteria weights
528.5	E: So, I reckon it's a four or five.	Suggests a formal weighting for the criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally
529	B: Oh yeah.	Agrees with above suggestion on the importance weighting of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally
529.5	C: I'd say five.	Suggests an importance weighting for criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally
530	D: I think it's five.	Agrees with above suggestion on the importance weighting of criterion 'force required'	Determine importance of criterion; Criterion 'force required'	Weighting criteria formally

Evaluate alternative-criterion pair

Below are various transcript excerpts showing activity patterns related to the process step 2B.2A/B.3. They are all taken from the transcript of the professionals' decision-making process. The different shades indicate that the examples represent different constructs (of the same type, see section 3.4.6 and figure 3.4).

First excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
21.03	C: 'Efficiency of delivery', 'time of each delivery'?	Controls the evaluation process by suggesting the consideration of another criterion	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Controlling the process
21.04	B: You have got to set the thing up ...	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
21.08	C: It's immediately a 'negative' isn't it?	Evaluates concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Evaluating restricted performances formally
21.15	C: If these guys have a job to finish, everything that adds a few minutes to each delivery is gonna be a negative.	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
21.19	A: But once it is set-up it depends on ... how many cylinders you gonna get delivered.	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
21.27	B: But, you still have to ... each time he's got to go down to the bottom to ...	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
21.31	C: ... unload it.	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
21.32	B: It's because it's a one-man thing whereas this (the current) is a two-man thing ... so it's a slightly unfair comparison.	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
21.43	A: Doesn't the landlord become involved on this for positioning them? (addresses the researcher)	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining the product environment
21.49	Researcher: Actually, the landlord shouldn't be asked for any help.	Provides external information on the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Gaining external information
21.52	C: Right, so the landlord ...	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining the product environment
21.54	A: ... is not in the equation.	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining the product environment
21.56	C: So, single man operation.	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining the product environment
22.08	B: Each time he delivers something he is gonna go up and down the stairs. He goes downstairs and puts them (the cylinder) where it belongs and he goes upstairs and gets the next one.	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
22.18	A: But, here he is lowering three at a time ...	Raises evidence for the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
22.21	B: Em, true ... so it is three times as efficient as ...	Deliberates the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally
22.24	C: yeah	Agrees with the above deliberations on the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally
22.25	B: Put a minus.	Evaluates concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Evaluating restricted performances formally
22.27	C: I think it takes longer ...	Deliberates the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally
22.41	C: Yeah.	Agrees with the above deliberations on the performance of concept A in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally
22.43	B: Anyway, what's the next one?	Controls the evaluation process by suggesting to move further in the evaluation process.	Evaluate alternative-criterion pair; Concept A; Criterion 'time of each delivery'	Controlling the process

Second excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
27.54	B: Can it deal with different cellars?	Controls evaluation process by suggesting a particular evaluation	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Controlling the process
27.56	A: Yes.	Deliberates the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Determining or evaluating restricted performances informally
27.58	B: Hmm, well again ... no ... I would say not as good as a bloke.	Deliberates the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Determining or evaluating restricted performances informally

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
28.09	C: I mean we are comparing it with the current method of delivery ... and the current method of delivery is a man turning up doing the job!!!	Raises evidence for the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
28.24	B: Can it cope with different sorts of cellars ...	Controls evaluation process by reminding the group of the comparison basis	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Controlling the process
28.25	C: ... in the same way as a delivery man?	Controls evaluation process by reminding the group of the comparison basis	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Controlling the process
28.26	B: Probably not.	Deliberates the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Determining or evaluating restricted performances informally
28.29	B: I mean it wouldn't be able to cope with the cellar we visited.	Raises evidence for the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
28.31	C: Oh no.	Agrees with the above evidence for the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
28.32	A: No chance.	Agrees with the above evidence for the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
28.34	C: But a fellow could. He has to ... they do, they find a way round, don't they?	Raises evidence for the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
28.37	A: Yeah.	Agrees with above evidence	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
28.47	C: So for 'different cellars' we are saying this is a minus, yeah?	Evaluates concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Evaluating restricted performances formally
28.52	B: Yeah.	Agrees with the above evaluation	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Evaluating restricted performances formally
28.55	C: You (person A) are not, are you?	Controls the evaluation process by prompting a team member to speak	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Controlling the process

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
28.58	A: I hear what you are saying, but ... I think that could cope with other cellars, not that particular one that we went to ... but we have got to look at it generically, haven't we? Okay, you gonna have a few odd ones, but let's look at it ... the big percentage ... the higher percentage it will just be able to cope with it.	Determines the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Determining or evaluating restricted performances informally
29.16	B: Put equals.	Evaluates concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Evaluating restricted performances formally
29.18	C: Yeah, go equals, yeah.	Agrees with above evaluation	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Evaluating restricted performances formally
29.23	C: 'Environmental'?	Controls the evaluation process by suggesting the consideration of another criterion	Evaluate alternative-criterion pair; Concept A; Criterion 'environmental'	Controlling the process
29.25	B: Well, hang on ... just ... go back ... if it didn't have a barrel ramp ... if it was just one of these ones where you drop the barrel just straight down onto some sacks ... you wouldn't be able to install that there ... it's got to be installed like that ... you need to have some sort of access to the cellar like that ... and if you only had the vertical access then you would have problems installing that.	Raises evidence on the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
29.53	A: True.	Agrees with above evidence	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
29.55	C: But we are assuming ...	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Making assumptions on the product environment
29.57	A: ... that is just a small percentage.	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Making assumptions on the product environment
29.58	C: ... that's the extreme end of it.	Deliberates the product environment	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Making assumptions on the product environment

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
30	B: Alright.	Agrees with the above assumption	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Making assumptions on the product environment
30.04	A: I mean, in that particular case, it's another parameter, isn't it? Then you could maybe add a clip onto that design: In this situation put this (points on drawing) on, like a little hoist or something like that ... clip it on.	Raises evidence for the performance of concept A in respect of criterion 'different cellars'	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts
30.2	B: Right, okay.	Agrees with above evidence	Evaluate alternative-criterion pair; Concept A; Criterion 'different cellars'	Raising evidence on restricted performance of concepts

Third excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
35.18	C: Okay, and 'weight to be handled'?	Controls the evaluation process by suggesting the consideration of another criterion	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Controlling the process
35.22	A: Marginally more, ...	Determines the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Determining or evaluating restricted performances informally
35.26	A: ... because of the carrier ...	Raises evidence for the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Raising evidence on restricted performance of concepts
35.3	C: If we were assuming that we are saving time, because we are delivering 3 at the same time ...	Raises evidence for the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Raising evidence on restricted performance of concepts
35.35	C: ... it could be less weight to be handled.	Determines the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Determining or evaluating restricted performances informally
35.4	B: It doesn't have to be (speaks to person A) ... it's (the carrier) got the potential to be quite light...	Raises evidence for the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Raising evidence on restricted performance of concepts
35.45	B: ... and it's got to be better than the previous one ...	Evaluates the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Determining or evaluating restricted performances informally

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
35.51	C: Oh, it's better than the previous one, but it's more than the current.	Evaluates the performance of concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Determining or evaluating *restricted* performances informally
35.55	B: Well, maybe make that one (concept A) a double.	Evaluates concept A in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept A; Criterion 'weight to be handled'	Evaluating restricted performances formally
35.57	A: If it's done manually with a rope it's gonna be more, if it's done with a mechanism it's gonna be less.	Deliberates the sub-issue 'how to operate the carrier' and its influence on the current evaluation pair	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Discussing sub-issues
36.04	B: I'm not sure that's what it means.	Deliberates the meaning of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Defining criteria
36.1	C: We were assuming the 'weight to be handled' includes the apparatus and the cylinders.	Deliberates the meaning of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Defining criteria
36.15	B: Make the previous one a double negative and make this one a negative.	Evaluates concept A and concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concepts A and B; Criterion 'weight to be handled'	Evaluating restricted performances formally
36.21	C: We are assuming it will be by rope then!	Deliberates the sub-issue 'how to operate the carrier' and its influence on the current evaluation pair	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	*Accepting* assumptions on sub-issue solutions
36.24	A: No, you couldn't do that, it would be too heavy.	Deliberates the sub-issue 'how to operate the carrier' and its influence on the current evaluation pair	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Discussing sub-issues
36.27	C: Okay, in that case ...	Deliberates the sub-issue 'how to operate the carrier' and its influence on the current evaluation pair	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Accepting assumptions on sub-issue solutions
36.3	C: ... It becomes what? Equal or just a negative, single negative?	Controls the evaluation process by urging the group to come to a conclusion of the above discussion	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Controlling the process
36.37	A: Single negative.	Evaluates concept B in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concept B; Criterion 'weight to be handled'	Evaluating restricted performances formally

Fourth excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
43.14	C: Em, 'time of each delivery'?	Controls the evaluation process by suggesting the consideration of another criterion	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Controlling the process
43.25	C: Right, considering ... the last one we had has a positive because we said it would take less time ... with this, if you just set this up and you are able to shoot the ...	Raises evidence for the performance of concept C1 in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
43.34	A: I don't think you are able to let it free fall, though.	Deliberates sub-issue 'how to control the cylinders'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
43.36	C: I don't believe you should.	Deliberates sub-issue 'how to control the cylinders'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
43.38	A: No, you can't, because you have got the problem then of getting the cylinders back up. So, you can have some form of attachment device on there, or connection or whatever.	Deliberates sub-issue 'how to control the cylinders'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
43.46	C: That's why it's saying here 'pulled up by rope'.	Explains working principle of concept C1	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Clarifying conept working principles
43.49	A: Hmm.	Deliberates working principle of concept C1	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Clarifying conept working principles
43.5	C: Not necessarily lowered by rope but pulled up by rope.	Explains working principle of concept C1	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Clarifying conept working principles
43.52	A: But you would have to have something down there. I mean, you could then let it go once you have tied a rope to it.	Deliberates sub-issue 'how to control the cylinders'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
44	C: So what are we saying?	Controls the evaluation process by prompting the group to conclude the argument	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Controlling the process
44.02	B: It's no worse than the previous one and we said the previous one was better ...	Deliberates the performance of concept C1 in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
44.05	C: Yeah.	Agrees with the above deliberations	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally
44.06	B: ... providing you can say: do it ... so you can put three cylinders down and then go downstairs and take three cylinders out.	Raises evidence for an evaluation of concept C1	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
44.14	C: Which I ... I have a problem with that. That looks as if it is a single shoot.	Raises evidence for an evaluation of concept C1	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
44.21	A: That is a single shoot.	Raises evidence for an evaluation of concept C1	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
44.22	B: Yeah, but ... you could have a double ... there is nothing about that says it would be impossible to have a double bowled one.	Deliberates sub-issue 'moving multiple cylinders with C1'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
44.31	C: Hmm... you in fact load them... you had 3 in a row ... you just pull one out, the others come down afterwards.	Deliberates sub-issue 'moving multiple cylinders with C1'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
44.43	B: Hmm, you might find that ...	Deliberates sub-issue 'moving multiple cylinders with C1'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
44.46	A: You've got to be careful, because there is safety involved on that ... if you break the neck of that and it suddenly goes ... you have got a valve, haven't you ... it's got a plastic head on it ... you are causing accidents.	Deliberates sub-issue 'moving multiple cylinders with C1' in respect of criterion 'safety'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
44.59	C: So what's the assumption?	Controls the evaluation process by trying to come to a conclusion with the sub-issue 'moving multiple cylinders with C1'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Controlling the process
45.01	B: Well, you are not, because if you are doing it properly there is nobody in the cellar ... of course, if you're cheating and you've got the landlord there, unloading it ...	Deliberates sub-issue 'moving multiple cylinders with C1' in respect of criterion 'safety for staff'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
45.08	C: Yeah, but if you are down there yourself, and you are pulling one out and the others come down ...	Deliberates sub-issue 'moving multiple cylinders with C1' in respect of criterion 'safety for operator'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
45.13	A: ... you will be asphyxiated.	Deliberates sub-issue 'moving multiple cylinders with C1' in respect of criterion 'safety for operator'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
45.17	C: Yeah ... well that's maybe something ... are we assuming ...	Deliberates sub-issue 'moving multiple cylinders with C1' in respect of criterion 'safety for operator'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Discussing sub-issues
45.19	B: Maybe we put equals ...	Evaluates concept C1 in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Evaluating restricted performances formally
45.21	B: ...because we have got a similar potential to put 2 down	Raises evidence for the above evaluation	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Raising evidence on restricted performance of concepts
45.24	C: Okay. (puts down a '=')	Agrees with the above evaluation	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Evaluating restricted performances formally
45.26	B: Put equals ...	Agrees with the above evaluation	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Evaluating restricted performances formally
45.28	B: ... because it is not quite as good as the previous one.	Evaluates the performance of concept C1 in respect of criterion 'time of each delivery'	Evaluate alternative-criterion pair; Concept C1; Criterion 'time of each delivery'	Determining or evaluating restricted performances informally

Fifth excerpt:

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
79.3	C: 'Weight to be handled'?	Controls the evaluation process by suggesting the consideration of another criterion	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Controlling the process
79.31	A: The same, isn't it?	Deliberates the performance of concept B and concept C2 in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Determining or evaluating restricted performances informally
79.32	C: The same.	Agrees with above deliberations	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Determining or evaluating restricted performances informally
79.33	A: Well no, B is more isn't it? ... B is a lot more.	Deliberates the performance of concept B and concept C2 in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Determining or evaluating restricted performances informally

Counter	Original conversation	Abstracted conversation	Aspect addressed	Activity
79.36	C: Does he do it ... ?	Deliberates the working principle of concept B	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Clarifying concept working principles
79.39	A: ... collectively!	Deliberates the working principle of concept B	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Clarifying concept working principles
79.4	C: Does he actually have to control the drop?	Deliberates the working principle of concept B	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Clarifying concept working principles
79.43	C: So, that's 2 and 4? (puts it down)	Evaluates concept B and concept C2 in respect of criterion 'weight to be handled'	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Evaluating restricted performances formally
79.46	A and B: Yeah.	Agree with above evaluation	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Evaluating restricted performances formally
79.5	B: Something we didn't consider was how much space it takes up on the lorry.	Identifies new criterion 'space requirement on lorry'	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Identifying criteria
79.55	C: No, we didn't.	Agrees with the above suggestion	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Identifying criteria
79.56	A: Well, the problem ... the problem starts here.	Discusses the meaning of criterion 'space requirement on lorry'	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Identifying criteria
79.59	B: I just think one of the criteria ... if the equipment fills up half of the lorry then you wouldn't ... you would be carrying less ...	Discusses the meaning of criterion 'space requirement on lorry'	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Defining criteria
80.06	C: Yeah ...	Agrees with above discussion	Evaluate alternative-criterion pair; Concepts B and C2; Criterion 'weight to be handled'	Defining criteria