Assessing the Importance of Factors Determining Decision-Making by Actors Involved in Innovation Processes

Hans Heerkens

Innovations can be seen as chains of non-routine decisions. With each decision, the innovator has to assess how important the various decision attributes are. Because the decisions are non-routine, innovators cannot fall back on judgements of past importance. Most decision support methods elicit importance judgements but do not help innovators or other decision-makers with the mental processes leading to the judgment. The 'importance assessment process' can be divided into seven phases (such as (sub-)attribute processing and various forms of weighting). The phase '(sub)-attribute processing' is the most important phase in terms of effort devoted to it, and the most obvious pitfalls that prevent valid importance assessments appear in this phase. This article describes some of these pitfalls. A few simple instruments may provide better-founded importance judgements that can be better communicated to other actors involved in innovation processes.

Introduction

This article is about a specific aspect of decision-making: the assessing of the importance of attributes (characteristics that describe the alternatives from which the choice is to be made). If you want to buy a new car, the cars available on the market (the alternatives) can be described in terms of their top speed, their price, their roominess and other attributes. Which car you buy not only depends on the score on these attributes (for example: how fast can car X go?), but also on their importance; sometimes referred to as their weights. If top speed is important to you, you may buy that expensive and cramped Ferrari. But if you have to take the kids to school each day, roominess may be more important than top speed, and you will buy a Volkswagen. In this article we address the question 'how do people think when assessing importance of attributes?'. This thinking process is called 'importance assessment', and it results in an 'importance judgement': a weight assigned to each attribute relevant for the decision at hand. The decision process is visualized in Figure 1.

In this article, we are only concerned with step two: the importance assessment.

We focus on the decision context that innovators often find themselves in, and show which pitfalls innovators face and how these can be negotiated. The most important phase in the importance assessment process, (sub)-attribute processing, is the focus of this article because it offers the most possibilities for improvement. We start with characterizing the situation in which an innovator often finds himself, and explore the problems he faces concerning the weighting of attributes. Then we review earlier research, culminating in a model of the importance assessment process. Our own research identified some major pitfalls in the importance assessment process and as a conclusion of this article we provide suggestions for some simple instruments that may be of help to those involved in innovative decision processes.

Note: when 'he' or 'his' is used, 'he/she' or 'his/her' is meant.

The Problem

An innovator is often confronted with an array of decisions to take; choices to make. For example: which technical solution is the best

doi:10.1111/j.1467-8691.2006.00401.x

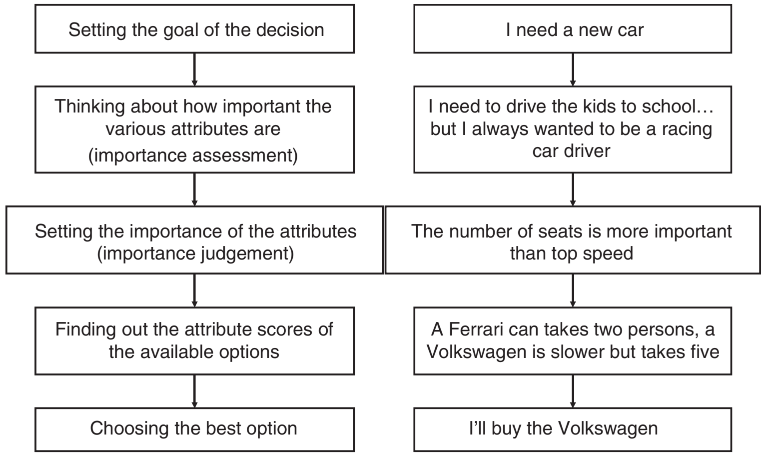

Figure 1. The Place of Importance Assessment in the Decision Process

for a given problem? Which of his colleagues in the organization should he let in on his ideas in order to gain support? Is it worthwhile to apply for a patent? Every choice has its pros and cons. Sometimes these are clear, like the advantages and disadvantages of technical solutions or the cost of patents and the legal protection they can give. In this article, we assume that the pros and cons pertaining to decisions are known. But even so, the relative importance of these pros and cons may not be known, either because they are a matter of personal preferences or because the situation in which the innovator finds himself does not have a precedent, so that past experience is of only limited value. Imagine, for example, that someone has invented a new model skateboard, and that type-A wheels give much less friction but are more prone to breaking off than wheels of type B. Which wheel should the inventor choose? This depends on how important performance (dependent on friction) is, relative to safety (if a wheel breaks off the skateboarder may be injured). In addition, replacing a wheel will incur extra cost. How important are these costs? With a new product, market research (finding out how important customers consider the various attributes to be) is difficult because the customers may not be able to give valid opinions. The innovator has to rely on his own judgement.

There is one simple reason for the difficulty in judging the importance of attributes; the attributes cannot be directly compared to each other. So the innovator has to rely on his expert knowledge and experience to assess the importance of the various pros and cons of the available options (in our example the types of skateboard wheels). The chance that there is a standard procedure that guarantees an optimal decision is negligible.

All in all, innovators frequently face the task of assessing the importance of attributes that influence the choices that they have to make. This article is about the way people do this, and how they can do it better. We focus on one phase of the importance assessment process, as stated earlier, and on the challenges that innovators face in that phase. The problem is non-routine, no clear criteria exist for judging which solution is the best, and evading a choice or leaving it to someone else is not an option. Innovators, faced with choices for which no precedents exist, are inevitably confronted with the importance judgement question. Many will rely on intuition, experience or 'gut feeling'. They may make the right choices, or they may not. In short, innovators:

- are faced with non-routine decision, or 'wicked problems' (Rittel & Webber, 1973);
- often have to take crucial decisions by themselves, unable to rely on importance assessments made by others. Even if they can ask others to evaluate their choices, they first have to make them themselves. Their importance assessment is, at least initially, theirs alone;
- have a strong need to justify their decisions in order to overcome opposition to their ideas;
- are sailing uncharted waters in which even their own expertise is limited.

So, innovators often operate in a different environment from decision-makers confronted with more routine tasks. The latter can fall back on their own experience with similar decisions, can call in help from others and probably do not even have to make explicit importance judgements. They can use importance judgements made in earlier decisions. The innovator does not have this luxury, and weighting is a truly challenging task. His confidence could receive a major boost if he could understand the mental processes involved in importance assessment, and if simple but effective instruments could be developed to help with weighting.

Two limitations of this article should be stressed. First, it is focused on individual actors and not on group decisions. Many business decisions are, at least, prepared by groups and not individuals. But innovators are often individualists. Once a decision is placed on the agenda of a designated group the innovator apparently managed to get it there, so in a sense his work is already done. But even in group decisions the participants have to decide individually on their position as input for the group process. And, even though a group may propose a choice, it may be up to an individual actor to make it. Our skateboard inventor may have obtained the advice of engineers and marketers on the configuration of his board, but as chief designer he has to make the final decision.

The second limitation is that the creative element of importance assessment is not addressed directly. We will see that creativity certainly plays a role in importance assessment process, especially in the phase that we will look at, but space constraints prevent us from going into this area.

Theoretical Background

Previous Research

An important area of research in decision theory linked to importance assessment is the study of factors that influence the assessment process. Notable examples are: perception of risk and attitudes towards risk (Kahneman & Tversky's (2000) Prospect Theory), the perspective of the decision-maker (Kray, 2000; Kray & Gonzalez, 1999), information presentation and usage (Guo, 2001; Russo, Medvec & Meloy, 1996), the concept of attribute weights (Keeney & Raiffa, 1976), factors that cause biases in weighting (Borcherding, Schmeer & Weber, 1995; Fischer, 1995; Póyhónen & Hámáláinen, 1998), the influence of unimportant or irrelevant attributes on choice (Barlas, 2003; Goldstein & Busemeyer, 1992; Hsee, 1995), group decision-making, particularly the relationship between individual and collective preferences (for example Hollingshead, 1996; Wei et al., 2000), the internal and predictive validity of various methods for measuring importance judgements, such as conjoint measurement (Harte & Koele, 1995; Jaccard, Brinberg & Ackerman, 1986) and the influence of regret aversion (Zeelenberg et al., 1996). Marketing research has focused on socio-economic and demographic factors influencing the perceived importance of attributes, and Meehl (1954) studied the degree to which limited numbers of relevant illness symptoms occurring with patients influence doctors' diagnosis. Further contributions regarding the explanatory power of regression models have been realized by Dawes (1979) and others. Keeney's (1992, 1994) value-focused thinking approach shows how various instruments assessing value preferences of decision-makers can be used to optimize non-routine decisions within an organizational context.

The above-mentioned research provides us with some building blocks for modelling importance assessment processes, such as the concept of 'importance' and the relationship between attribute scores, weights and attractiveness of alternatives. But it either takes the weights actors assign for granted and concentrates on the decisions made on the basis of these weights, is concerned with eliciting the weights with sufficient validity (necessary for linking them with choices) or looks at the factors influencing the weights, such as perceptions of risk. The importance assessment process, that is, the way in which actors think while weighting, is addressed merely incidentally. So, our research covers the process of generating weights, while the above-mentioned research concerns the weights when they are, or have been, assigned and measured with methods such as conjoint measurement.

There are other areas of research that are of relevance to importance assessment, such as problem-solving, human choice strategies and bounded rationality (Heerkens, 2003). We will devote no attention to them here, but instead go straight to the model that we developed of the importance assessment that actors go through in non-routine decisions. This model is explained in detail in Heerkens (2003) and Heerkens and Van der Heijden (2003). From the model we can derive the aspects that are of particular interest to innovators. This model was made for exactly the situation in which innovators find themselves, and which was described earlier. We will select specific issues from the model that are relevant for innovators. We do not aim to discuss the model or the research that generated it in depth; that is done elsewhere (Heerkens & Van der Heijden, 2003, 2005).

The Weight Assessment Model (WAM)

The weight assessment model (WAM) consists of seven main phases and six auxiliary activities. The seven main phases are presented in a

Table 1. The phases of the WAM

Phase	Phase name	Percentage of segments devoted to the phase
1. Structuring cluster	Problem identification	6.74
2. Structuring cluster	(Sub-) attribute processing	30.33
3. Weighting cluster	Absolute sub-attribute weighting	27.22
4. Weighting cluster	Homogeneous sub-attribute weighting	4.53
5. Weighting cluster	Heterogeneous sub-attribute weighting	1.50
6. Weighting cluster	Attribute weighting	12.54
7. Evaluation cluster	Evaluation	17.14

sequential way in Table 1. In reality, actors may go back and forth between phases and are likely to address phases more than once.

The auxiliary activities pertain to areas such as information search and planning and are not linked to particular phases. We do not deal with them in this article, but a brief description can be found in Appendix 1. Based on Simon's (1960) distinction of the problem-solving process into a structuring and a solving phase, we divide the WAM in a structuring cluster (phases 1 and 2) and a weighting (solving) cluster (phases 3 to 6). Phase 7, the evaluation phase, will not be covered in this article. We confine ourselves to the weighting itself, not to its evaluation. In the structuring cluster, the problem (in this case the task to weigh attributes) is formulated (phase 1) and the attributes are processed in phase 2 so that they can be readily weighted in subsequent phases. Then, the weighting takes place in the weighting cluster.

Since the model was conceived on the basis of think-aloud protocols, we are able to indicate which percentage of all the thoughts that the subjects expressed was devoted to which phase. We arrived at these percentages by dividing each protocol in segments (the smallest possible meaningful statements made by the subjects) and coding these according to a formal coding scheme. The total numbers of statements pertaining to each of the phases was divided by the gross total of statements for all subjects taken together. This is further explained in Heerkens and Van der Heijden (2003, 2005). We use these percentages as an indicator of the relative importance of each phase.

Phases of the WAM

Phase 1: Problem Identification

This phase consists of activities such as elaborating on (understanding, concretizing) the task at hand and, if desired, re-formulating it in one's own words. Essentially, this phase concerns defining, so to speak, the task lying ahead. This may mean, for example, stating the attributes to be weighted, or making boundary conditions explicit. In our research, the subjects had to weigh safety and comfort of a minibus (explained later in this article). This means that other potentially relevant attributes, such as fuel consumption, were to be ignored. The generation of attributes to be weighted may take place in this phase, but it is also possible that they were identified before and merely have to be made explicit.

Phase 2: (Sub-)Attribute Processing

If one wants to weigh attributes, one should first know what one is weighting. Attribute-processing concerns giving the attributes a more precise meaning. This can be seen as a case of framing (Akin, 1994). An example of attribute processing is: concretizing 'safety of a minibus' into 'number of deaths per 10 million km'.

Attribute properties such as measuring level, measuring unit, level of abstractness and precision can change as a result of processing. The following forms of processing were identified:

- Decomposing. An attribute ('safety') can be split up in several sub-attributes (such as 'quality of the brakes' and seatbelts present yes/no);
- Re-formulating. When an actor gives an attribute or sub-attribute a different name while meaning the same attribute with a similar, not necessarily identical, measurement unit, the attribute is re-formulated. For example, 'safety' can be re-formulated as 'ensuring a safe journey';
- Concretizing a (sub-)attribute. An example was given above;

- Integrating (sub-)attributes into a new (sub)-attribute. For example: taking 'adjustability of seats' and 'amount of legroom' together as 'quality of seating';
- Making an attribute more abstract. This is the complement of concretizing.

Definition is not included in the types of processing. The result of processing may be a description of an attribute that is so exact and formal that it can be called a definition. This is the phase that we will study in some depth in this article. The next phases concern the actual weighting process.

Phase 3: Absolute (Sub)-Attribute Weighting

With 'absolute' weighting (based on Timmermans, 1993), we mean that a statement about the importance of a (sub)-attribute is made without reference to the importance of other (sub)-attributes. For example: 'safety is important'. This statement does not say how much more or less important it is than 'comfort'.

Phase 4: Homogeneous Sub-Attribute Weighting

This phase is the first in which 'true' weighting takes place: the balancing of the weight of one sub-attribute against that of another. We call this 'relative weighting' (based on Timmermans, 1993). In this phase, two or more sub-attributes of the same main attribute are weighted against each other, and arguments for the weighting are given. For example: 'good seatbelts are more important than good brakes' (both sub-attributes of 'safety').

Phase 5: Heterogeneous Sub-Attribute Weighting

This phase differs in only one respect from the previous one: the sub-attributes that are weighted belong to different main attributes. For example: good seatbelts (sub-attribute of 'safety') are more important than comfortable seats (sub-attribute of 'comfort').

Phase 6: Attribute Weighting

This phase concerned the integral weighting of the (in our case, two) main attributes. It is the essence of any weighting task. For example: safety is more important than comfort'.

Phase 7: Evaluation

This phase comprises the reflections by subjects on their activities and the results. Several types of evaluation can be identified, such as the extent to which the assignment has been fulfilled, evaluations of weight judgements (is the actor, on hindsight, satisfied with assigned weights) and evaluations of arguments (how good are the reasons for particular weight judgements).

From this model, several areas can be identified that are of particular significance to innovators, because they present special challenges or potential pitfalls. In this article we focus on phase 2; the main structuring phase. If attributes are not correctly formulated, the resulting weighting is bound to be flawed. In addition, it is clear from the table that phase 2 is the most important phase in terms of effort devoted to it (30 percent of the segments in the think-aloud protocols pertain to this phase), which seems logical in the case of non-routine problems. Furthermore, when an innovator has to weigh attributes, it is almost inconceivable that the attributes are so clear that no processing is needed. After all, an innovator has to deal with ill-defined, non-routine problems. So, if we want to help innovators to improve their weighting of attributes in the difficult decisions they have to take, phase 2 seems to be a good place to start. (Sub-)attribute processing has, to our knowledge, not been addressed in the literature (the WAM was, after all, only designed a few years ago), so it is worthwhile to explore it.

In the remainder of this article, we will address the following questions:

1. How do actors involved in weighting attributes in non-routine decisions conduct the attribute processing phase?
2. Which pitfalls can be identified that actors should be aware of?
3. In which way can their performance be improved?

Obviously, we have to answer question 1 and 2 before we can hope to answer question 3.

The Research Method

In order to gain insight in cognitive processes of our subjects, we used a think-aloud method. This is a good method for analysing cognitive processes (Ericsson & Simon, 1993). Methods such as choice experiments and process tracing show the results of cognitive processes, but not the processes themselves, while retrospective reporting methods, such as interviews and diaries, leave too much room for interpretation of the cognitive processes by the subjects themselves and are vulnerable to lapses of memory (Ericsson & Simon, 1993).

As there is little research on importance assessment processes (see above), we do not formulate elaborate hypothesis, but we describe the attribute processing phase on a number of dimensions and link these to existing theory whenever possible. As stated before, we will not address creativity. It is clear that creativity probably plays an important role in generating (sub-)attributes, especially in non-routine decisions where there are no established sets of (sub-)attributes available. In addition, finding integrative attributes covering sets of sub-attributes can be a creative activity. We will, however, look only at the types of processing actors employ, not at their source of inspiration. We will look at pitfalls in the process from a rational perspective, not at enhancing creative activities. We do this because of space constraints and in order to keep our research focused.

Sample and Assignment

Eighteen undergraduate students of the University of Twente in The Netherlands were given an individual assignment based on a fictional case. University students might be assumed to have enough analytical abilities to perform the assignment satisfactorily, without having the knowledge and skills that would enable them to rely on previous experience of importance assessments. Hence, the danger that they give weights based on previously obtained knowledge is minimized.

The assignment consisted of supporting the acquisition process of new minibuses by a local company. This decision may not be innovative, but it certainly is non-routine for a local (small) company. Such a decision occurs only every so many years, and choosing the wrong bus can have dire consequences (for example: low load factor if the bus is too big, loss of clients if it is uncomfortable and so on). The management and the drivers may have practical experience with the minibuses the company operates, but they may well not have intimate knowledge of the buses available on the market, and of the demands of (potential) clients. A non-routine decision like this is a good case for studying the importance assessment process.

The subjects were asked to establish the importance of two characteristics of the to-be-acquired minibuses *vis-à-vis* each other while thinking aloud, and were told that they would be advising the management team during the acquisition process. The management team would use the generated weights as a basis for evaluating the available minibuses, and then choose the type to buy. The management team did not have to agree on the weights generated, but the students should be able to explain their reasons behind their importance judgements if so required by the management. Therefore, while the students were free in the way to reach their importance judgment, it was not a purely personal exercise but an activity within an organizational context. The attributes, ‘safety’ and ‘passenger comfort’, were chosen to prevent comparability by some readily available algorithm or heuristic or easy expression in a common denominator such as money. The information that was supplied included a brochure of the company, a leaflet explaining the decision context and two brochures on minibuses; one on a Volkswagen and one on an Opel. The latter enabled the subjects to get familiar with the specific capital good to be acquired. It was made clear that these examples of minibuses did not mean that the subjects had to make a choice between them. The students were given an example of weighting before starting the assignment, but were allowed to use their own concept of ‘weight’ or ‘importance’, just as would be the case in real life.

It should be stressed that the assignment was geared to provide the optimal context for an importance assessment process. This means that there was no direct relationship with innovation. The context was kept as simple as possible, so as not to distract the subjects from their task. A purchasing context is much easier to grasp than an innovative context, where our subjects could not have been expected to generate the innovation themselves anyway. The relevant circumstances, however, were there: a non-routine assignment where the subjects could not rely on previous knowledge or experience and had to use their own creativity and cunning to complete the task. It does not matter whether the importance assessment had to be done for an acquisition process (like that in our assignment) or for, for example, the judgement of innovative alternative solutions to a problem (generated by someone else).

Procedure

The respondents were asked to think aloud during the assessment process. The general guidelines for think-aloud studies given by Ericsson and Simon (1993) were followed, including a practice session to familiarize the subjects with the think-aloud strategy. All verbal information given by the respondent was recorded and typed out literally. After completion of the assignment, a short interview was conducted. In total, each session lasted for a maximum of two hours, for which the subjects were paid €20. Two pilot sessions were conducted, which led to some minor adjustments of the assignment.

Two kinds of analyses have been performed using the literally typed out protocols:

1. A largely qualitative analysis according to the general rules of the 'grounded theory' approach (Strauss & Corbin, 1998).
2. A more quantitative analysis, based on a formal coding scheme that was designed on the basis of the qualitative analysis. Two coders performed the coding activities. Although they worked independently of each other, during the coding of the first six protocols weekly meetings were held to discuss general coding issues in order to enhance the reliability of it. The coders retrospectively applied the refinements to the coding scheme independently. The overall Cohen's Kappa (Baarda & de Goede, 2001) for inter-rater consistence was 0.97 over a total number of verbal segments of 7,253.

Limitations

It should be stressed that this is not quantitative research. It was not clear beforehand which variables would turn out to be relevant and how they should be defined. Hence we used an extreme case and no control variables. We wanted to know how actors having to make an importance assessment for a non-routine problem would behave and which pitfalls they could encounter: whether, for example, a more routine problem would induce a different behaviour was not of interest. If we can identify pitfalls that innovators have to be aware of because they have a reasonable chance of occurring (even if the exact chance is not known) or because innovators can easily recognize them in their individual activities, regardless of the chance of occurring, our aim is fulfilled. If we could have linked the occurrence of pitfalls, or the behaviour of actors making importance assessments, to the quality of the resulting weighting, then it would have been appropriate to control that behaviour or the occurrence of the pitfalls. But we were unable to define a valid measure of 'quality of the weighting' that is usable in a laboratory context. An 'extreme case' in which we gave the subjects maximum freedom as to how to fulfill the assignment, so that we could expect a maximum of variation in behaviour, could be expected to yield more useful, although less precise, results than an experiment with a limited set of well-defined (control) variables.

This research was conducted with a small group of subjects because the in-depth analysis of the think-aloud protocols was very time-consuming. As far as the qualitative analysis was concerned, it turned out that, after 12 or so protocols had been analysed, little new insights came from each extra protocol. So we accepted the lack of statistical validity as the price to be paid for in-depth qualitative analysis. The subjects were students with no prior experience with either the acquisition of minibuses or formalized organizational importance assessment processes. The research was conducted in a laboratory context. This means that the results have limited statistical and external validity. The first is, in our view, not a great problem. The trends in the results seem to be quite clear and multiple indicators were used for many variables, thereby increasing internal validity. It also, however, means that no definite conclusions can be drawn for other groups from the group that we studied. We can, and do, make propositions about how actors in real-life situations may behave, based on our results and on the literature. The basic regularities in importance processes that we describe will, we expect, be present in some form in real-life situations. After all, it is not uncommon for individuals (albeit often with a certain degree of expertise) to make importance assessment processes under circumstances similar to those in our research. So our research provides a basis from which to look at real-life situations: inventors, professionals and policymakers. Our research pertains only to decision contexts where there is explicit weighting and where the importance assessment process is separated from the evaluation of alternatives.

Results: How Actors Conduct the (Sub-)Attribute Phase

Earlier, we mentioned five forms of attribute processing: decomposing, concretizing, integration, making abstract and re-formulating. We will use these to describe our results.

Decomposing Attributes

Decomposing attributes can have the important function of making clearer what is actually meant with an attribute, and also whether it can be used to make a complex attribute measurable. 'Safety' for example, is a very broad concept. It can be seen as a potential outcome of action (the chance of dying in a traffic accident), as an attribute that increases the chances of a desired outcome (good brakes, so that accidents can be avoided), as an emotion (feeling safe) and so on. Of course, if one wants to attach importance to safety, one has to know what one means by it. It also may turn out that 'safety' is actually a collection of sub-attributes, each with their own weight. The weights of the sub-attributes together should add up to the weight of 'safety' (Keeney & Raiffa, 1976), but chances are that they do not always do so in practice, although experimental findings are as yet inconclusive (Borcherding, Schmeer & Weber, 1995; Fischer, 1995; Póyhónen & Hámáláinen, 1998) Even if the actual weighting of attributes is only rudimentary, as it often will be when an innovator weights attributes (since his ideas are only in development and not yet crystallized), decomposing can give some anchors. It enables (sub-) attributes to be compared as to importance at an ordinal level, it shows the consequences of giving certain (sub-)attributes higher or lower weights and generally helps to 'paint a picture' of attributes in plotting a course of thought or action. So, decomposing can be functional during importance assessments. Our research shows, however, that it is used often in such a way as to be dysfunctional:

- so many sub-attributes are generated that oversight is sure to be lost;
- decomposing is not done systematically;
- nothing is done with many of the sub-attributes that are the result of decomposition.

Table 2 shows the number of sub-attributes for 'safety' at various levels of decomposition. A level is defined as the number of splits that resulted in a certain sub-attribute. So, if 'comfort' is split in a number of sub-attributes, amongst which 'quality of the seats', which in turn is split in 'width of the seats' and 'height of the armrests', then there are two levels of decomposition. For 'comfort', the numbers were roughly similar to those of 'safety'. It can be seen that the attribute is decomposed in a large number of sub-attributes. The average number of sub-attributes per subject for 'safety' was 19.6. For comfort the number was 24.4. Fifteen subjects generated ten or more sub-attributes for safety. All subjects generated ten or more sub-attributes for comfort.

It is not difficult to imagine that one quickly loses oversight with such numbers of attributes. Short-term memory is limited to 7–10 items (Miller, 1956). As most subjects did not write down the sub-attributes, they would find it very difficult to work with them. The problem was compounded by the fact that decomposition was purely associative; in fact almost completely unsystematic. For example: 'safety' can very well be split into 'active safety' (aimed at preventing accidents by, say, good brakes) and 'passive safety' (mitigating the consequences of accidents, as seatbelts are designed to do). From this point, one could go on generating sub-attributes for both types of 'safety'. But not one subject worked like this. If divisions like 'active' and 'passive safety' were used, then they were put next to, instead of above, other sub-attributes. An example of one subject's attribute processing is given in Appendix 2.

An excellent use of decomposing would be to establish causal, or at least statistical, relationships between (sub-)attributes. An example of such a scheme is given in Figure 2.

A scheme of causal relationships serves to eliminate unnecessary attributes and avoids double–counting (see, for example, Vincke, 1992). For example: some subjects thought that high weight makes a minibus unsafe, as does a long stopping distance. But braking distance is a function of, amongst others, weight (according to the formula acceleration is mass times force). So, 'weight' is superfluous and can be left out. That is, unless it has other effects on safety than through stopping distance, for example because high weight means a strong

Table 2. The Decomposition of 'Safety'

Number of attributes as a result of decomposition	Number (%) of subjects, first level	Number (%) of subjects, second level	Number (%) of subjects, third level
0		4 (22)	15 (83)
1–5	2 (11)	7 (39)	1 (6)
6–10	5 (28)	4 (22)	1 (6)
11–15	5 (28)	2 (11)	1 (6)
16–20	5 (28)	1 (6)	
21–25			
26–30			
30–35	1 (6)		

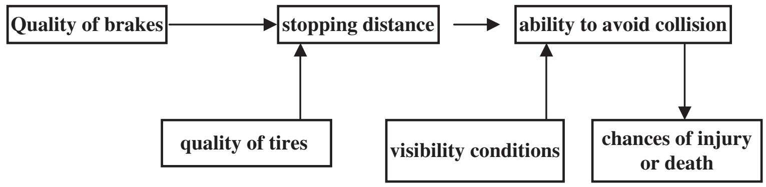

Figure 2. A Scheme of Causal Relationships Between Sub-Attributes

Table 3. Integration of Safety and Comfort

Number of sub-attributes being integrated	Number (%) of subjects integrating sub-attributes of safety	Number (%) of subjects integrating sub-attributes of comfort
0	9 (50)	13 (72)
1–5	6 (33)	5 (28)
6–10	3 (17)	

chassis. But then we should leave either 'weight' out, or 'stopping distance' and 'strength of the chassis'. If we do not, we count 'weight' twice, once directly and once through its effects. However, no subject even tried, much less succeeded, in establishing causal relationships. They sometimes mentioned causal relationships in passing, but they did not make any systematic use of them.

Going one step further, a possible use of establishing causal relationships between (sub-)attributes is: find a common denominator for expressing some or all attributes to be weighted. With this, the weighting problem is effectively eliminated. For example, if it were possible to link both 'safety' and 'comfort' to 'profit' and it were turn out that 'profit' would be maximized with 10 units of 'safety' and 15 units of 'comfort', then the optimal decision is clear. No further weighting is required. This approach is difficult, but it is used in, for example, the acquisition of civil and military aircraft. Attributes are causally linked to, respectively, life-cycle cost and revenues and combat effectiveness. Yet none of the subjects even tried to find a common denominator. Only one of the subjects mentioned this as an explicit aim during certain stages of the assignment, but he did not pursue this aim consistently. The others did not address the quest for a common denominator in any systematic way.

We will return to the causal scheme when we suggest ways of improving the importance assessment process.

Knowing the causal relationships between attributes makes integrating attributes possible; the opposite of decomposition. With luck, all attributes can be integrated in two or three main attributes that can then be causally related to a common denominator. At the very least, integration reduces the number of sub-attributes to be weighted and makes the weighting less complex. How did the subjects go about integrating?

Integration of Sub-Attributes

There was hardly any integration of sub-attributes, contrary to what one expects when weights are to be given not to sub-attributes but to main attributes. Integration was rare, as can be seen from Table 3. This table shows the number of (sub-)attributes integrated. We classify a subject as a non-incidental user of integration if integration occurs in at least four cases over safety and comfort together. We see that at least half of the subjects do not integrate at all, and the other half integrates only a fraction of the sub-attributes generated.

Another way to look at the significance of integration is to observe how many (sub-)attributes are the result of integration. The maximum number of attributes that were the result of integration was four (one subject); two integrated attributes were found with only three subjects. Only in two instances was a sub-attribute resulting from integration given a weight during the final weight assignment. Integration always resulted in a new sub-attribute, not in the main attributes to be weighted according to the assignment (safety and comfort). The logic of the integration was often implicit and nearly always purely qualitative. No indexing or other quantitative methods were used. In sum, integration was by and large irrelevant.

That leaves only re-formulation, abstraction and concretizing to be addressed. We can be short about these. Their roles in attribute processing were minor, both quantitatively (number of abstractions and so on) and qualitatively (contribution to the weighting process). Only in five cases were between five and ten sub-attributes of either safety or comfort concretized, and never more than ten. Nine subjects in all made either safety or comfort concrete, but then they either proceeded with decomposition or eventually weighted only safety and comfort and not the sub-attributes. Only half of the subjects made abstractions for safety and for comfort, and never for more than two sub-attributes. We will not devote any attention to abstraction. Sometimes (sub-)attributes were reformulated, but this was almost exclusively limited to inconsequential changes such as 'comfort' becoming travelling comfort'. Whenever reformulations were more significant they could be classified as concretizations or abstractions.

Is it a good or a bad thing that there is so little concretization? At first sight, a logical function of concretization is operationalization. One subject used concretization for this purpose, and made it explicitly known that he wanted to operationalize sub-attributes and did so by taking the judgement of outside experts (for example, the Dutch Consumer Association) as an indicator of safety, driving quality, and some other attributes.

However, this subject was the only one who used concretization in this way explicitly and with any pretence of being systematic. So, although subjects used concretization to make (sub)-attributes more concrete, they had no use for the end-product of specification other then getting a better idea of the meaning of the (sub)-attributes. A way of concretization that was used by a number of subjects was explicitly or implicitly specifying the extremes of an attribute. This amounts to the beginning of scale construction. Three subjects (17 percent) used this directly in order to define weights. One subject used two types of cars to define the extremes of comfort (a Limousine and a Ford Fiesta). He used them only once, as examples, so this did not have any further measurable influence on the execution of the assignment.

Subjects did not evaluate the result of specification. It seemed to be a largely unintentional process, even more so than decomposition, where subjects sometimes used some system, however unsophisticated.

Discussion: Pitfalls in (Sub-)Attribute Processing

The attribute-processing phase is meant to convert attributes in such a way that they can be weighted. This can be done by defining them so clearly that their importance can be fixed, or by finding a common denominator. Our subjects, for the following identified reasons, achieved neither aim:

- there was no system in the decomposition;
- no causal relations were established between sub-attributes so;
- integration, concretization, abstraction and re-formulation hardly took place and;
- the unworkably huge number of sub-attributes was thus not reduced.

All this makes it unlikely that assessing the weights of the sub-attributes would in any way be easier than assessing the weights of the main attributes. It could be assumed that subjects used the processing of attributes mainly for framing purposes, that is, to find out what 'safety' and 'comfort' actually mean. But since all subjects devoted considerable effort to phase 3 (absolute weighting, mainly of sub-attributes, to be covered in a forthcoming article) the processing of sub-attributes appeared to mean more than just a framing function. The subjects obviously were not merely interested in obtaining concepts of 'safety' and 'comfort'. The sub-attributes were important in their own right. But it is unclear what their relevance is.

What does all this mean for an innovator in real life, wrestling with a choice to make between several alternative courses of action, each with their own pros and cons? It means a high degree of uncertainty in deciding what is important and what is not. The sub-attributes in the innovator's mind, and hence the assigned weights, may vary from one moment to the next. When there are a lot of (sub-) attributes, such variations can alter choices significantly. It also becomes more difficult to give good and consistent arguments for a decision or weight judgement, and to communicate those arguments to others. This is especially likely with 'wicked problems', the type of problems that an innovator is confronted with almost by definition. It is clear that ambiguity does not help in convincing others, or in securing alliances. There is ample research to show that people's weighting is often inconsistent with logic, in that choices made are not in line with the weights assigned (see, for example, Kahneman, 1994; Kahneman, Knetch & Thaler, 1990, 1991). They will sometimes choose from two options the one that is *least* attractive, given the weights they chose and the scores on the attributes. For example, they may choose the attribute with the highest score on the most important attribute, ignoring the alternative with a much higher score on a marginally less important attribute that makes it, all-in, the most attractive. And it gets worse. Levine, Halberstadt and Goldtsone (1996) showed that people's weighting gets less consistent when they have to deliberate about the weights. That may not be surprising, in the light of our research. It is conceivable that if people do not have to reason, they do not consider many sub-attributes of the attributes to be weighted, so their image of the attributes stays reasonably constant over time. Once they get to reason about the attributes, the images start changing as various sub-attributes are activated sequentially. Unfortunately, for an innovator having to choose a path of action, not reasoning about weights is not an option. So there is good reason to look for ways to improve the consistency and hence the quality of the weighting process.

Let Us Make Things Better: Improving Performance of (Sub-)Attribute Processing

Our advice for anyone who has to weigh attributes in an non-routine decision is simply to make the (sub-)attribute processing phase of the importance assessment process more transparent. This can be done in the following ways:

1. Instead of splitting attributes, try defining (concretizing) them. For example: 'safety' can be described as 'the number of deaths and wounded per 1,000,000 km where attributes of the minibus are the root cause'. With 'safety' thus specified, it may no longer be necessary to split it up in sub-attributes, because the decision-maker does not have to know which attributes play a role in accidents. The manufacturer of the bus knows, and gives a score for 'safety'.
2. If you feel you have to decompose, do it systematically. Some decision-makers may not be satisfied with (1) and they may want to specify an attribute further by splitting it up in sub-attributes. But, then do not just start throwing sub-attributes around. For 'safety', several possible decomposition systems are possible: active versus passive safety, attributes concerning the engine, brakes, chassis, lighting or interior and so on. Such a system makes it easier to assess the completeness of the list of sub-attributes (to be covered in a future article) and makes the next two points easier to perform.
3. Make a 'cognitive map' (de Boer, 1998; Warren, 1995) of the causal relationships between (sub-)attributes. A cognitive map is simply a drawing of the attributes connected by arrows. The arrows represent cause-effect relationships. Figure 2 is essentially an example of a cognitive map. In this way, the really important attributes stand out, as do the superfluous ones (remember the example of weight, stopping distance and strength of the chassis). If you are lucky, you end up with only four or five relevant sub-attributes per attribute. Those can all be taken into account if the importance of the attribute is to be set, much more so than the 19 to 25 sub-attributes our subjects generated per attribute.
4. Try to integrate sub-attributes to a level that is as high as possible but at which you feel you can still assign weights. For example: it may not be more difficult to set a weight for 'active safety' than it is to weigh 'quality of the brakes' or 'having power steering'. Here again; the fewer (sub-)attributes to weigh, the better. Fewer (sub-)attributes means more consistency (hopefully), a better oversight of what you are doing (certainly) and a better argumentation to give if your decision is challenged (very likely).
5. After having done one or more of the previous points, decide at which level the (sub-)attributes should be weighted. As we showed in Table 1, our subjects frequently generated three levels of sub-attributes. Even after having left out and/or integrated some sub-attributes, more than one level may be left. It is not a foregone conclusion that weighting has to be done at the lowest level. Sure, it promises more precision, but there are more sub-attributes to weigh and, ideally, all sub-attributes should be weighted against all other sub-attributes to get the optimal result (for a method to do this, see de Boer, 1998; Saaty, 1980). This can be a daunting task.

Closing Remarks

We have shown that the most important phase of the importance assessment process, phase 2 of the WAM, contains pitfalls that actors can not only identify but can also negotiate. Hopefully, although we cannot prove it, improving the quality of the (sub-attribute processing phase will improve the quality of the subsequent weighting phases. And it may increase the confidence of the actors concerned in their weighting, which they may experience as a positive effect in itself. The instruments discussed above are conceptually simple and can be utilized with only a sheet of paper at hand and a few hours' time to spare. They are ideal for individual actors with few resources at their disposal and the need to clarify for themselves, if not for others, the pros and cons of their decisions. Thus, they can be considered ideal for innovators (who often, although not always, work alone in conceptualizing their new ideas) to carry them in their mental arsenal. But the use of these instruments has to be learned. In the experience of the author with under- and postgraduate university students, making a cognitive map, for example, takes at least an afternoon to practise. A global understanding of the elements of a decision (alternatives, attributes and weights) is also needed. In addition, it is not clear how much the entire importance assessment process, and the decisions based on it, will improve by focusing on the (sub-)attribute phase But it seems the results are well worth the limited training effort required.

References

Akin, Ó. (1994) Creativity in Design. *Performance Improvement Quarterly*, 7, 9–23.

Baarda, D.B. and de Goede, M.P.M. (2001) *Basisboek methoden en technieken*. Stenfert Kroese, Groningen.

Barlas, S. (2003) When Choices Give in to Temptations: Explaining the Disagreement Among Importance Measures. *Organizational Behavior and Human Decision Processes*, 91, 310–21.

Boer, L. de (1998) *Operations Research in Support of Purchasing; Design of a Toolbox for Supplier Selection*. Universiteitsdrukkerij, Enschede.

Borcherding, K., Schmeer, S. and Weber, M. (1995) Biases in Multi-Attribute Weight Elicitation. In Caverni, J.P., Bar-Hillel, M. and Jungemann, H. (eds), *Contributions to Decision Making-1*. Elsevier Science BV, Amsterdam, pp. 3–28.

Dawes, R.M. (1979) The Robust Beauty of Improper Linear Models in Decision Making. *American Psychologist*, 34, 571–82.

Ericsson, K.A. and Simon, H.A. (1993) *Protocol Analysis: Verbal Reports as Data* (2nd edn). Bradford Books/MIT Press, Cambridge MA.

Fischer, G.W. (1995) Range Sensitivity of Attribute Weights in Multiattribute Value Models. *Organizational Behavior and Human Decision Processes*, 62, 252–66.

Goldstein, W.M. and Busemeyer, J.R. (1992) The Effect of Irrelevant Variables on Decision-Making: Criterion Shifts in Preferential Choice. *Organizational Behavior and Human Decision Processes*, 52, 425–54.

Guo, C. (2001) A Review on Consumer External Search: Amount and Determinants. *Journal of Business and Psychology*, 15, 505–19.

Harte, J.M. and Koele, P. (1995) A Comparison of Different Methods for the Elicitation of Attribute Weights: Structural Modeling, Process Tracing and Self-Reports. *Organizational Behavior and Human Decision Processes*, 64, 49–64.

Heerkens, J.M.G. (2003) Modeling Importance Assessment Processes in Non-Routine Decision Problems. Enschede (dissertation).

Heerkens, J.M.G. and Van der Heijden, B.I.J.M. (2003) On Importance Assessment and Expertise in Non-Routine Decisions; an Exploratory Study on the Cognition of Weighting Processes of Capital Goods' Attributes. *International Journal of Management and Decision Making*, 3, 370–98.

Heerkens, H. and Van der Heijden, B.I.J.M. (2005) On a Tool Analyzing Cognitive Processes using Exploratory Think-Aloud Experiments. *International Journal of Human Resources Development and Management*, 5(3), 240–83.

Hollingshead, A.B. (1996) The Rank-Order Effect in Group Decision-Making, *Organization Behaviour and Human Decision Processes*, 68, 181–93.

Hsee, C.K. (1995) Elastic Justification: How Tempting but Task Irrelevant Factors Influence Decisions. *Organizational Behavior and Human Decision Processes*, 62, 330–7.

Jaccard, J., Brinberg, D. and Ackerman, L.J. (1986) Assessing Attribute Importance. *Journal of Consumer Research*, 12, 463–7.

Kahneman, D. (1994) New Challenges to Rationality Assumption. *Journal of Institutional and Theoretical Economics*, 150, 18–36.

Kahneman, D., Knetch, J.L. and Thaler, R.H. (1990) Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, 96, 1325–48.

Kahneman, D., Knetch, J.L. and Thaler, R.H. (1991) Anomalies: the Endowment Effect Loss Aversion and Status Quo Bias. *Journal of Economic Perspectives*, 5, 193–206.

Kahneman, D. and Tversky, A. (ed.) (2000) *Choices, Values and Frames*. Cambridge University Press, Cambridge.

Keeney, R.L. (1992) *Value-Focused Thinking*. Harvard University Press, Cambridge MA.

Keeney, R.L. (1994) Creativity in Decision Making with Value Focused Thinking. *Sloan Management Review*, 35, 33–41.

Keeney, R.L. and Raiffa, H. (1976) *Decisions with Multiple Objectives, Preferences and Value Tradeoffs*. John Wiley & Sons, New York.

Kray, L.J. (2000) Contingent Weighting in Self-Other Decision Making. *Organizational Behavior and Human Decision Processes*, 83, 82–106.

Kray, L. and Gonzalez, R. (1999) Differential Weighting in Choice versus Advice: I'll Do This, You'll Do That. *Journal of Behavioral Decision Making*, 12, 207–17.

Levine, G.M., Halberstadt, J.B. and Goldtsone, R. (1996) Reasoning and the Weighting of Attributes in Attribute Judgments. *Journal of Personality and Social Psychology*, 70, 230–40.

Meehl, P.E. (1954) *Clinical Versus Statistical Prediction*. University of Minnesota Press, Minneapolis.

Miller, G.A. (1956) The Magical Number Seven, Plus or Minus Two. Some Limits on Capacity of

Processing Information. *Psychological Review*, 63, 81–7.

Póyhónen, M. and Hámáláinen, R. (1998) Notes on the Weighting Biases in Value Trees. *Journal of Behavioral Decision Making*, 11, 139–50.

Rittel, H.W.J. and Webber, M.M. (1973) Dilemmas in a General Theory of Planning. *Policy Sciences*, 4, 155–69.

Russo, J.E., Medvec, V.H. and Meloy, M.G. (1996) The Distortion of Information During Decisions. *Organizational Behavior and Human Decision Processes*, 66, 102–10.

Saaty, T.L. (1980) *The Analytic Hierarchy Process*. McGraw-Hill, New York.

Simon, H.A. (1960) *The New Science of Management Decision*. Prentice-Hall, New Jersey.

Strauss, A. and Corbin, J. (1998) *Basics of Qualitative Research; Grounded Theory, Procedures and Techniques*. Sage Publications, Thousand Oaks CA.

Timmermans, D. (1993) The Impact of Task Complexity on Information Use in Multi-Attribute Decision Making. *Journal of Behavioral Decision Making*, 6, 95–111.

Vincke, P. (1992) *Multiple Decision Aid*. John Wiley & Sons, Chicester.

Warren, K. (1995) Exploring Competitive Factors using Cognitive Mapping. *Long Range Planning*, 28(5), 10–21.

Wei, Q., Yan, H., Ma, J. and Fan, Z. (2000) A Compromise Weight for Multi-Criteria Group Decision Making with Individual Preference. *Journal of the Operational Research*, 51, 625–34.

Zeelenberg, M., Beattie, J., van der Pligt, J. and de Vries, N.K. (1996) Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making. *Organizational Behavior and Human Decision Processes*, 65, 148–58.

Hans Heerkens (j.m.g.hererkens@utwente.nl) is assistant professor at the School of Business, Public Administration and Technology, department of Operational Methods for Production & Logistics, University of Twente, PO Box 217, 7500 AE Enschede, Netherlands, phone (–31)-53-4893492

Appendix 1: The Auxiliary Activities

During the formal coding of the protocols, some activities were found that could not be placed in particular phases, because they occurred in several phases or because they logically are not part of the weighting process or to the assignment. The first category comprises activity planning and information assessment. Examples of the second category are alternative judging (choosing a minibus) and attribute scoring (assessing whether a minibus is, for example, comfortable). These activities certainly are phases of the acquisition process but they should follow the weighting process, not be part of it. Also in the second category is weighting procedure design for a real-world situation. The assignment did not call for this and the designed procedures could not be used in the experimental situation. Expressing emotions is part of moth categories. The table shows the auxiliary activities of the WAM.

Table A1. Auxiliary Activities of the WAM

Activity number	Activity name
1	Alternative judging
2	Attribute scoring
3	Activity planning
4	Information assessment
5	Weighting procedure design
6	Expressing emotions

We will now briefly discuss the auxiliary activities.

1 Alternative Judging

Subjects can make judgements about the attractiveness of alternative minibuses (mainly the two minibuses mentioned in the information package) or about the attractiveness of minibuses in general vis-à-vis other modes of transport, such as trains or private cars.

2 Attribute Scoring

Subjects might make statements about how they think a particular minibus, minibuses in general or alternative modes of transport might score on certain (sub)-attributes. Examples of such statements are: 'Both the Opel and the Volkswagen have a stereo-set on board'.

3 Activity Planning

Prior to starting one of the phases of the WAM, a subject might plan how to execute the phase. An example is the statement: 'I think I should start by defining what I think "safety" and "comfort" actually mean'.

4 Information Assessment

With information assessment, we mean activities concerned with searching for information, and assessing the value of the available information for the task at hand. All subjects, at one point or another, concerned themselves with issues like the sort of information that they felt was needed, the information that could or could not be found in the information package and the quality of the information provided.

5 Weighting Procedure Design

Since the subjects were instructed to imagine that they had to perform the weighting in support of an acquisition process of an organization, they might indicate how they would propose to conduct the weighting process (or the acquisition process as a whole) if they were really working for a company and not just imagining it. Some subjects, for example, stated that they would in reality propose to conduct a market survey in order to assess how important safety and comfort are to present or potential customers.

6 Expressing Emotions

Statements such as: 'Oh, how difficult this is! I don't think that I can do it', and: 'I am distracted by all the birds I see flying outside' perhaps do not say very much about the actual weighting process, but they form a clearly identifiable category and they might affect the weighting. For example, if a subject feels he or she cannot cope with the assignment and still has this feeling after weights have been given, the confidence level probably is very low, which could lead to equal weights for all attributes.

Appendix 2: Example of An Attribute-Processing Scheme by One of the Subjects (Students) in Our Research

The schemes should be read as follows. Safety always gets the number 1 and comfort the number 2. Decomposed attributes at the first level get the numbers 1.1, 1.2, 2.1, 2.2 etc. At the second level, the numbers consist of three digits and can be, for example, 1.1.1, 1.1.2 etc. A letter placed after a certain attribute number means that the attribute is a re-formulation. If an attribute is the abstraction of another attribute, this is noted between brackets. An integration is always the result of two or more attributes being processed, and is also indicated between brackets.

The sub-attributes of the first level are listed as much as possible in the order in which the subjects mentioned them.

This is the (sub-)attribute processing scheme of just one of the subjects in our research, of only one of the two attributes. The scheme is representative of the schemes made by the other subjects, in terms of size and structure.

The Processing of 'safety'

1. *Safety*
   - 1a: If an accident happens, you want to get out in one piece, preferably unhurt (specification)
   - 1b: I want to get out in one piece or with very minor injuries, but not so that I can sit in a wheeled chair for the rest of my life
   - 1c: Accidents
   - 1.1: Number of deaths per year with a certain brand
   - 1.1a: Number of accidents with which it has occurred
   - 1.1b: Number of deaths per year with accidents (from the context it is clear that it is meant per type)
   - 1.1c: Accident numbers (from the context it is clear that it concerns deaths per year)
   - 1.1d: Maximum so many deaths per year
   - 1.1e: How many deaths per year with accidents and with how many accidents does this happen?
   - 1.1f: Number of deaths per year

1.2: Number of serious injuries
   - 1.2a: *Number of serious injuries per year per accident*
   - 1.2b: *How often does it occur (serious) injuries?*
   - 1.2c: *Figures about serious injuries*
   - 1.2d: *Number of serious injuries per year*
   - 1.2.1: *Paralyzed (downwards) from a certain body part or really loose a body part*
   - 1.2.1.1: *Paralyzed*
   - 1.2.1.2: *Body part coming off*

1.3: Seatbelts
   - 1.3a: *Are seatbelts in the car?*

1.4: Seat broke loose

1.5: Anti-skid system

1.6: Are there headrests?
   - 1.6a: *Headrests*

1.7: Can headrests be adapted?
   - 1.7a: *Are headrests adaptable?*
   - 1.7b: *Are they adjustable in height (no specification because this is what he meant with 1.7 and 1.7a)*
   - 1.7c: *Are headrests adjustable?*
   - 1.7d: *Adjustable headrests*

1.8: Safety for driver

1.9: Safety for assistant-driver

1.10: Safety for passengers

1.11: Airbag
   - 1.11.1: *Airbags on the side*

1.12: How does a bus fare if you smash into it from the front, the rear, the side and from above?
   - 1.12a: *With crash tests what was the result (abstraction)*
   - 1.12b: *Result with type of accident*
   - 1.12c: *Result with crash tests*
   - 1.12.1: *If an airplane crashes on your car*
   - 1.12.2: *If such traffic pole like you have in Enschede comes crashing into your car from underneath*
   - 1.12.3: *From the side they come*
   - 1.12.3a: *If someone comes from the side*

1.13: To what extent does a baby sit safely in the car?
   - 1.13.1: *Has it got baby seats?*
   - 1.13.2: *Does the possibility exist to install them (baby seats)*
   - 1/13.1/ 1.13.2(a) *Baby seats are they there, can they be installed (integration)*
   - 1.13.3: *Do baby seats have to be with the face forward or with the face rearward?*
   - 1.13.3a *Which baby seats are dangerous, which are not dangerous?*