INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'07
28 - 31 AUGUST 2007, CITE DES SCIENCES ET DE L'INDUSTRIE, PARIS, FRANCE

COMPARISON OF FOUR EXISTING CONCEPT SELECTION METHODS

Sauli Honkala[1], Matti Hämäläinen[1] and Mikko Salonen[1]
[1]Helsinki University of Technology, Finland

ABSTRACT

Successful companies need new products in a constant flow in today's highly competitive business environment. These products must stand out from competitors, the product quality must meet the customer demands and the profits must be adequate to keep stock-owners satisfied. Several important decisions are made at the early phase of each product development project. These decisions have an apparent impact on the final outcome of the project. For this reason, concept selection is an area in design science that is under considerable interest continuously.

Many criteria based evaluation methods have been proposed by academia in order to make the concept selection process more credible and transparent. These evaluation methods incorporate several different approaches and a variety of mathematical calculation models to assist deciding upon the best alternative or alternatives. In addition to the different approach, the time needed to conduct the evaluation varies according to the method used and the amount of comparison criteria and concept alternatives used.

The objective of this study is to compare existing multi-criteria decision making (MCDM) methods for concept selection, to identify possible differences in the methods, and to give recommendations for their use in concept selection in variable situations. In this paper, four MCDM methods were compared in eight cases that were adapted from various sources. The comparison primarily showed parallel results between compared methods, but certain noticeable differences also occurred. These differences are pointed out and clarified, and five suggestions for the general use of MCDM methods were made.

Keywords: Design methods, concept selection, multi-criteria decision making

1 INTRODUCTION AND OBJECTIVES

Concept selection is an activity in the product design process, where alternative concepts are compared and a decision is made to select the alternative(s) which proceed into the later phases of design [1]. Several authors have raised concept selection as one of the most critical issues in design [1, 2].

There are at least three remarkable challenges in concept selection. First, the nature of available information is usually based on subjective perceptions and speculations, and accurate calculations are seldom available [2]. Second, the stake-holders, users, designers and producers, can have conflicting requirements concerning e.g. product design and manufacturing, or product performance and sales price [2]. And third, the freezing of product concept can have far-reaching effects on product costs and customer satisfaction, which can only be fixed with additional costs and time [3].

Birkhofer studied the use of scientific methods in industry in general [4]. The conclusion of the study was that only very few of the scientific methods were actually used in industry. Birkhofer stated two reasons for this outcome. The methods used in industry lack the scientific background and are more or less based on experience rather than scientific testing. The second suggested reason was that the academia has not correctly understood the real industrial demand and application environment.

Salonen&Perttula [5] studied methods used for concept selection by Finnish industry, and found out that formal concept selection methods were seldom used in actual decision making. The methods used in reality were most often informal in nature. However, the companies that replied to have used formal methods were more satisfied in decision making process than the ones in which decision concerning concept selection were made informally.

Finding of Salonen&Perttula are in line with more general level findings of Birkhofer and they indicate that even if concept selection and MCDM-methods are widely studied by academia, the methods have not been broadly accepted and used in industry. These support that there still exist a need for further research in this subject.

The objective of this paper is to compare existing MCDM methods for concept selection, to identify possible differences in the methods, and to give recommendations for their use in concept selection in variable situations.

State-of-the-Art

Various multi-criteria decision making (MCDM) methods exists to aid concept selection. Each of these methods has its individual perspective into estimating the merit, rank and value of the compared concept alternatives. However, little research has been done to determine the differences between these methods, and the concept selection situations they are best suitable for. Although some results exist, e.g. Triantaphyllou has found that in certain situations different methods can give different answers to the same problem [6], the differences in the results of these methods have not been clearly clarified, nor impacts of these differences discussed.

Four commonly known MCDM methods are analyzed in this study. These methods are Pugh's concept selection (Pugh) [7], weighted rating method (WRM) as presented by Ulrich and Eppinger [1], Saaty's Analytical Hierarchy Process (AHP) [8], and Roy's Electre III (Electre) [9]. Weighted score of source values is used as a reference in comparisons. These methods were chosen because their use is widely spread and each of them is proposed to suite concept selection.

The selected MCDM methods vary from each other in two different ways. The major differences originate from the use of either a concept comparing approach, or a comparison based on a pre-defined rating scale. Pugh, WRM, and Electre use reference value or reference alternative in case of each criteria as a base for the rating [1,7,9]. AHP uses a pair-wise comparison method [8], where each alternative at a time is compared against the rest of the alternatives to calculate final score in each criterion.

In addition to difference in comparison method, also the rating scale varies between the methods. The Pugh method, where only ratings "–", "s", and "+" can be used for each comparison has the lowest rating resolution in compared methods [7]. The Electre and AHP methods represent the far-end of the methods in this case. In Electre, the rating scale can be freely selected in case of each criterion [9]. In case of AHP, rating scale with 17 steps from 1/9 to 9 can be used in each of the comparisons [8].

Concept selection methods can be divided into different categories in several ways. Takai&Ishii [10] classify these methods into perception based and analytical approaches. When analytical methods are used, a function describing product performance is established and the result is optimized by altering parameters of equation. Group of analytical methods was excluded form the scope of this study because these methods are usually developed to be product specific by using history data. Roy&Bouyssou [11] have presented another classification which is based on the calculation model of methods. They divide MCDM methods into total aggregation, hybrid, and partial aggregation methods. Pugh and WRM can be categorized as total aggregation methods, and Electre as a partial aggregation method. In this categorization AHP belongs to a group which is a hybrid of total and partial aggregation methods.

In this study, selecting methods with different calculation procedures was considered to be more interesting than comparing only methods within the same category. Therefore several methods, such as Fuzzy AHP [12], Quality Function Deployment [13] and Promethee [14], were left out of the comparison, since the selected ones are considered to represent their whole sub-group accurately enough.

2. RESEARCH METHOD

The comparison of the four MCDM methods was conducted by using the same source values that were derived from various individual cases. A universal scale of 1 to 100 was used for source values in each criterion. A total number of eight different types of concept selection cases were used in this study. Four cases were adopted from existing scientific publications, three cases were generated specifically for the comparison analysis to represent widely different concept selection situations, and one case was based on the known specifications of actual products on the market.

In each of the cases, several alternatives, i.e. product concepts, and several criteria exist. Since all of the evaluation methods have their own scale for evaluation, the source values needed to be converted appropriately. These conversion processes are discussed later. In order to compare the four MCDM methods, the method specific result values were scaled into a percentage (0–100 %) after the scoring process. The best alternative received a reference value of 100 percent, and the rest were given a value according to their relative score.

Our result analysis concentrated primarily on the scoring of alternatives and secondarily on the overall rank of each alternative. With these focuses we were looking for an answer to the following two questions:

1) By use of the MCDM methods, is the selection clear and can we identify one clearly best concept alternative?
2) If not, then which concept alternatives would we choose for further development based on the results of the MCDM methods?

Since e.g. Pugh and WRM methods use dissimilar algorithms to score and rank alternatives compared to AHP and Electre, our main hypothesis for the outcome of this research is that *different concept selection methods will give different results* in our analysis cases. However, we expect all methods to be capable of identifying the clearly best and clearly worst alternatives in those cases where such alternatives exist.

Source values

Eight concept comparison cases were used in this study. These comparisons and their references are described in table 1. The table also contains the case descriptions, and the number of concepts and criteria within each of the cases.

Table 1. Comparison case used in study

Case #	Case subjects	Source data format	No. of criteria	No. of alternatives	Reference or description
Data adopted from scientific publications					
1	Project selection	Financial estimate and 0..100 rating	5	5	Buchanan [15]
2	Contractor selection	AHP scale	6	5	Kamal [16]
3	Material selection	Material characteristics	7	5	Otto&Wood [17]
4	Product concept selection	AHP scale	4	4	Yeo [12]
Self generated cases					
5	Narrow margins	Rating range 1..100	5	8	Self generated
6	Strong differences	Rating range 1..100	5	8	Self generated
7	Normal distribution	Rating range 1..100	5	8	Random generated
Data adopted from product specifications					
8	LCD-monitor comparison	Product specifications	4	5	Product specifications

Comparison cases 1–4 are adopted from published scientific articles. Before the comparison could be done, their source data needed to be converted into a uniform scale.

Comparison five, “narrow margins”, represents a case, in which all of the concept alternatives have almost the same average score. Distribution of scores between concept alternatives varies strongly. Total of five alternatives and eight criteria were used. All of the criteria had the same weight factors. This case represented difficult decision making situation, and in this case it was assumed that the MCDM methods would produce conflicting results.

Comparison six, “strong differences”, represents a case, in which very strong differences exist between the alternatives. Total of five alternatives and eight evenly weighted criteria were used. By this we examined, how the methods would react to a strong difference between alternatives.

Comparison seven, “normal distribution”, used alternatives that had been produced by a random generator. Each of the performance characteristics, that were given values from 1 to 100, were generated by combining two random generated seeds between 0,5–50. Hence, the performance characteristics of concept alternatives were distributed between 1 and 100 according to normal distribution.

In comparison eight, “LCD-monitor comparison”, five LCD-monitors were compared. The technical data was collected from product specifications and was converted into universal scale. The appearances of the compared products were estimated subjectively.

Conversion between universal and method specific scales

All MCDM methods have a have a scale for rating the alternatives, and this scale is generally specific to the method in question. In our study, these scales are referred to as Method Specific Scales (MSS) in distinction from a general, 1–100 scale which is referred to as universal scale. In order to compare different methods, the universal values must be convertible to MSS and vice versa. Here we describe the formulas used in our comparisons to convert source values between MSS and universal scale.

In Pugh’s method, the MSS values used are “–“, “s” and “+” [7]. In this case, a percentage threshold value was used to define rating boundaries for each of the criteria. A fixed value of 50 in the universal scale was used as a reference for all of the criteria, and a threshold value of 20 % was used. Each of the source values were compared against these rating boundaries and appropriate MSS rating was given. For example, if an alternative was over 20 % better than the reference, a rating of “+” was given. If the alternative was over 20 % worse than the reference, the rating was “–”, and otherwise the rating was “s”.

The Weighted Rating Method uses MSS values from 1 to 5 [1]. In WRM, the rating is executed by comparing all of the concept alternatives against each other with respect to one criterion at a time. Ulrich and Eppinger [1] point out that it is not always appropriate to use a single reference concept for all of the criteria. Therefore, one criterion at a time, the concept alternative with a source value closest to 50 was set as the reference in current criterion and the rest of the alternatives were compared to this reference. Rating boundaries were calculated for each criterion separately with the following equations. Equation 1 was used to calculate upper boundaries for ratings 1 and 2, and equation 2 was used to calculate the upper boundaries for ratings 3 and 4. Figure 1 shows three examples of the rating boundaries. Universal scale rating is shown on the y-axis and the WRM rating on the x-axis.

$$Boundary(1,2) = [Val(ref) \cdot Y(Rating)] / 2,5 \qquad (1)$$

$$Boundary(3,4) = [100 - Val(ref)] \cdot [Y(Rating) - 2,5)] / 2,5 + Val(Ref) \qquad (2)$$

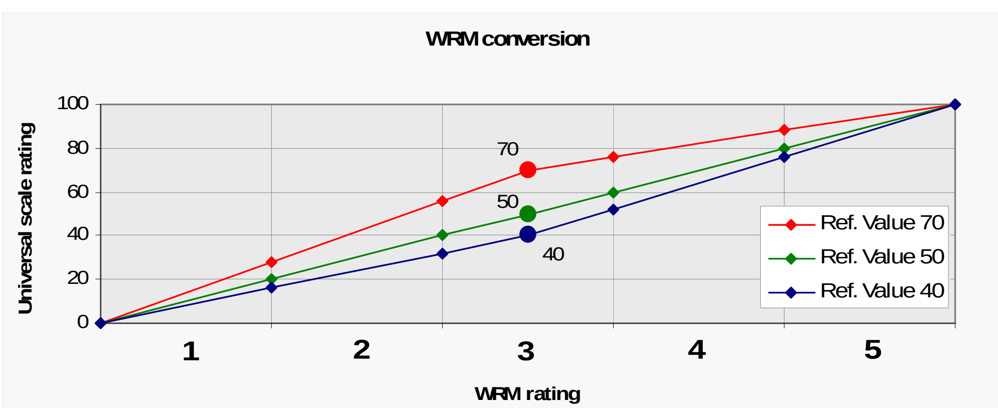

Figure 1.Conversion between Weighted Rating Method and universal scale

In Analytic Hierarchy Process, the MSS values are presented as fraction numbers from 1/9 to 1/2, and integers from 1 to 9 [8]. Conversion to AHP's MSS values contains several steps. Saaty points out that the scale is ratio-based [8], and therefore at first the source values needed to be converted to an A/B-ratio by dividing them with each other. Next, these ratio values were converted to the AHP scale.

An exponential function, as shown in figure 2, was used to define the boundaries for conversion. Each comparison ratio was compared to the rating boundaries, and a method specific value was given based on the corresponding AHP rating. After preliminary tests, the exponential function was adjusted so that a source value ratio A/B=5 produced a rating of 9 in AHP scale. This relation was selected in order to use the full scale of AHP, and each AHP matrix was checked afterwards with the consistency check [8].

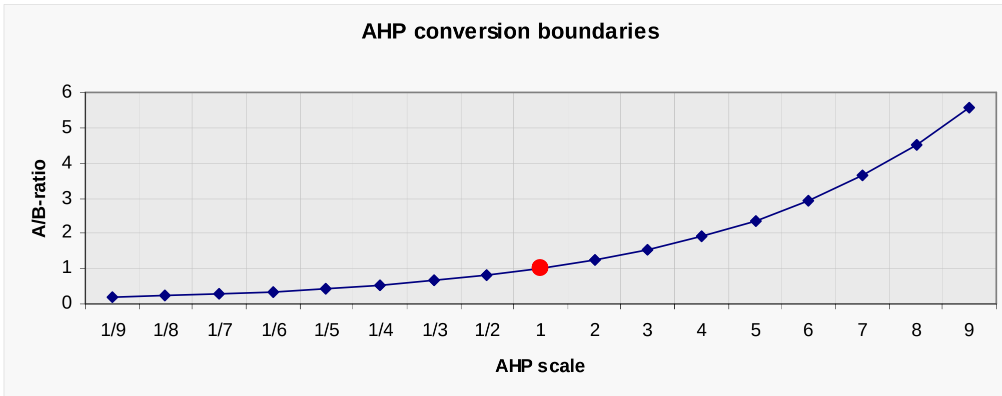

Figure 2. An exponential function was used to define the boundaries for conversion from a/b –ratio to AHP rating

The universal scale of 1 to 100 was used directly in Electre method. The threshold values were set to be 10 for the indifference threshold, 25 for the preference threshold, and 75 for the veto threshold, except in Buchanan's case [15], where the threshold values were directly adopted from the original case publication.

Source data from four scientific publications was used in this study. Already published and applied cases were used because these cases have already been proven to be suitable for this field of research. However, all of the data in different publications was presented in different, method specific scales and therefore had to be converted to universal values to be comparable. All of these conversions are explained in the following.

In Buchanan's project ranking study [15], four out of five criteria use values of 0–100. The fifth criterion is a financial estimate and is presented with an absolute value. In the first four criteria, zero values were replaced with value 1, because the AHP method requires source values to be above zero. In the fifth, financial criterion, the zero point was shifted and the range of values was scaled down in order to make values correspond to the universal scale.

The process of converting values from the universal scale to AHP's MSS, was used in reverse order to acquire data from Yeo [12] and Kamal [16] publications where data is presented in the AHP scale. Since the result of reversed conversion process is the A/B-ratio, some adjustments needed to be made. Ratio estimations were enhanced by taking into consideration that in the AHP method, the A/Bratio is in relation to ratios A/C and B/C etc. Finally, values A, B, C, etc. were scaled up so that none of the values were below 1 or above 100. Also averages in each of the criteria were taken into concern in the scaling. The achieved result represents only one possible solution and also other possible combinations can be found to produce the same result in the AHP method. The conversion process described above was tested and found to function accurately in several preliminary tests.

The material data acquired from Otto&Wood [17] was converted to be applicable by using the same equations as described in the paper with minor changes. Otto&Wood converted material data into the evaluation scale from -2 to 2 in their paper. The same equations were used in this study to convert the same material characteristics into ratings from -50 to 50. In order to this data to be applicable in our comparison, the mid-point was also shifted by 50 units.

3. RESULTS

In the results chapter, we present the eight cases examined in this study. The result focus is on the differences in ranking order and the overall score of each alternative in all of the cases. Based on our observations, the differences between methods are pointed out and their causes are identified.

Cases 2–6 proved out to be the most informative ones, and therefore they are presented first. Cases 1, 7, and 8 seem to be in line with our hypothesis, but since they have no additional value to the subject, their results are only presented as a summary.

Case 2: Kamal

In figure 3, according to WRM and Electre methods, alternative A3 should be selected and alternative A4 should be developed further if possible. The Pugh method identifies only alternative A3 to be selected. In deviation to others, AHP recommends selecting alternative four. The contradicting result of this particular case supports our main hypothesis.

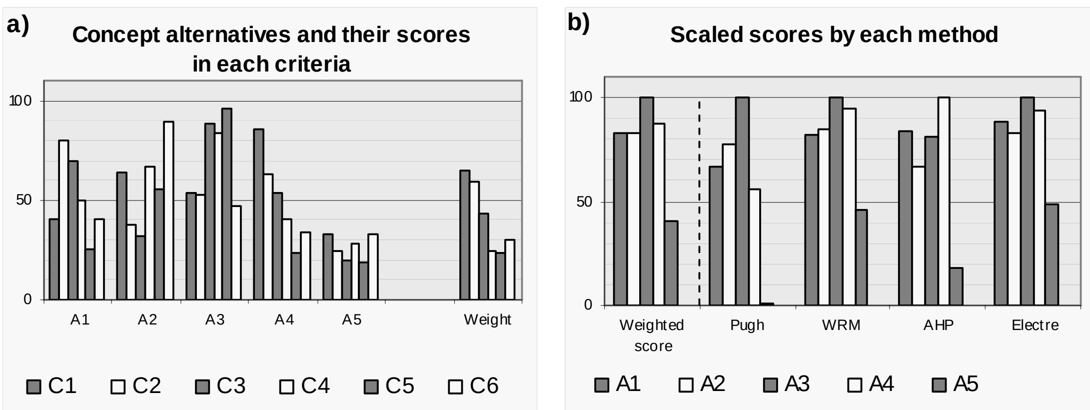

Figure 3. Concept evaluation of "Kamal, Application of the AHP in project management" – case. a) Source values in US, and b) Evaluation result

Case 3: Otto & Wood

In figure 4, according to the WRM, alternative A4 should be selected and alternatives A1, A2 and A5 should also be considered further. Pugh's method proposes alternative A2 to be selected but is unable to show any differences between rest of the alternatives. AHP proposes alternative A4 to be most suitable for selection and alternatives A1 and A2 to be developed further if possible. Electre method favors alternative A4, but the margin between the top scoring alternative A4, and alternatives A1, A2 and A5 is small.

In this particular case, Pugh's method recommended alternative A2, while the rest of the methods recommended alternative A4. However, the second position varies strongly between the compared methods. In the WRM case, the second in rank is alternative A4, in AHP case it is alternative A1, and in Electre case it is alternative A5. These differences support our main hypothesis.

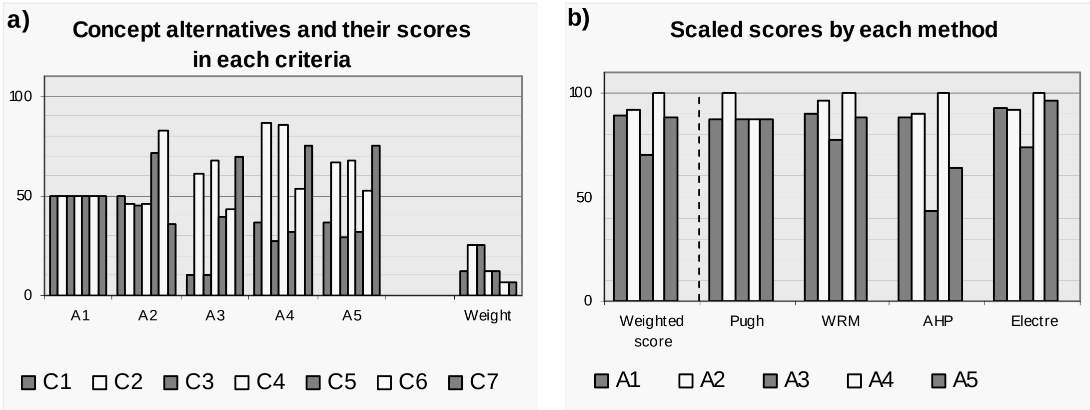

Figure 4. Concept evaluation of "Otto&Wood, Estimating errors in concept selection" – case. a) Source value distribution, and b) Evaluation result

Case 4: Yeo

The concept evaluation is relatively clear in figure 5. The ranking of alternatives is identical in all of the methods. The A4 alternative is the most promising one according to all of the methods. Only the Electre method proposes to develop another concept as well.

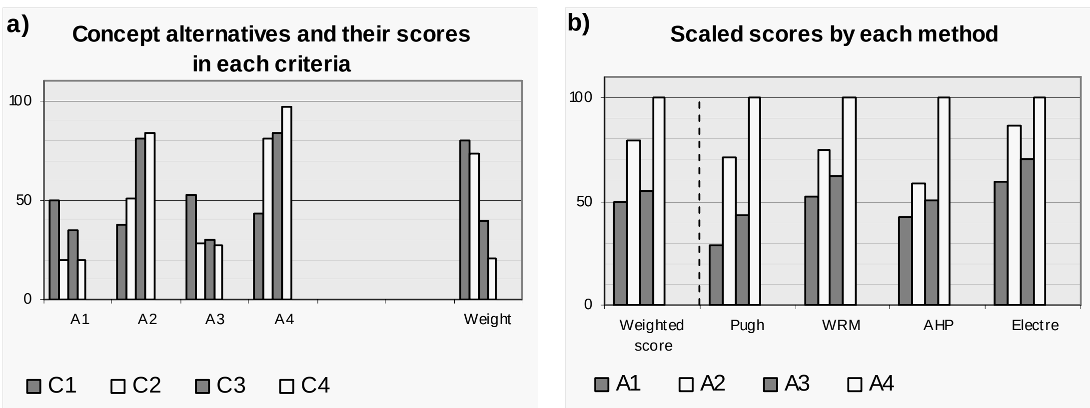

Figure 5. Concept evaluation of "Yeo, Analysis of decision-making methodologies" –case. a) Source value distribution, and b) Evaluation scores

Case 5: Narrow margins

The result of WRM and AHP is very similar in this case as shown in figure 6. The difference between WRM and AHP is in relative scores, where AHP shows larger margins. The ranking order in Electre is quite different from the others methods, even though the margins are very small.

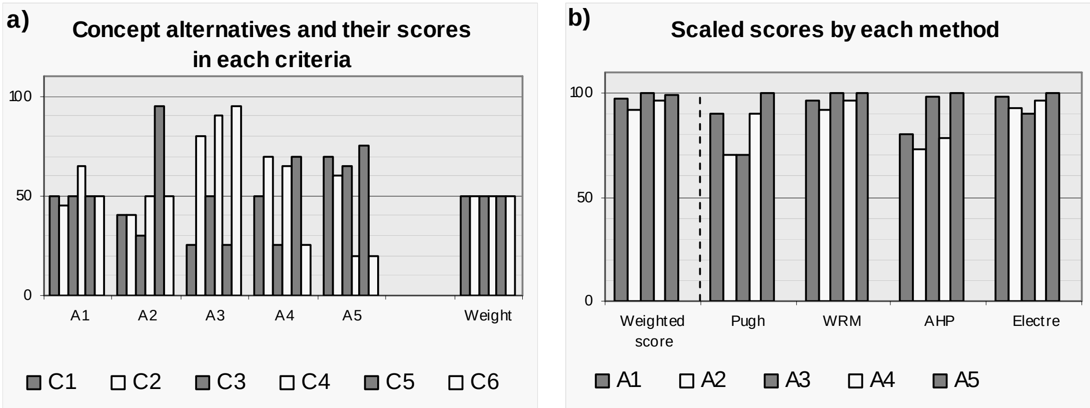

Figure 6. Concept evaluation of "narrow margins" –case.
a) Source value distribution, and b) Evaluation scores

Case 6: Strong differences

In figure 7, the overall rank of alternatives is identical in all methods except in Pugh. Even though, the ranking of alternatives is identical, large differences in relative scores can be found. This case supports our hypothesis that the methods are able to identify the clearly best and worst alternatives in the cases in which these appear.

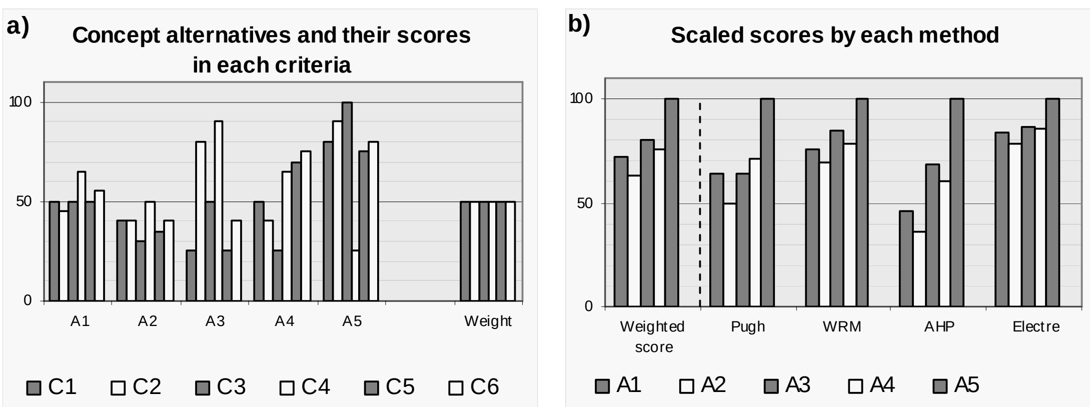

Figure 7. Concept evaluation of "Strong differences" –case.
a) Source value distribution, and b) Evaluation scores

Cases 1, 7 and 8

Cases 1, 7, and 8 are in line with our hypothesis, but no clear remarks can be made of them. All of these methods indicate the same alternatives as the best ones, and only small variations between the methods appear.

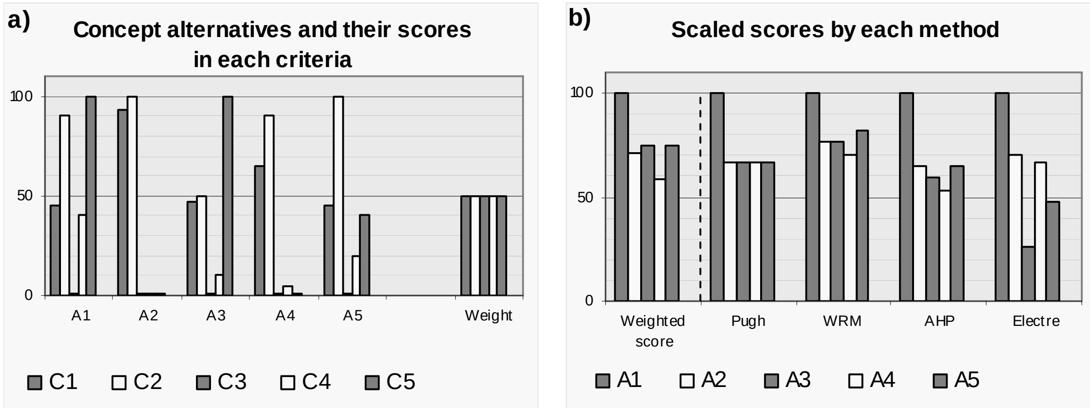

Figure 8.Concept evaluation of "Buchanan, Project ranking"-case
a) Source value distribution, and b) Evaluation scores

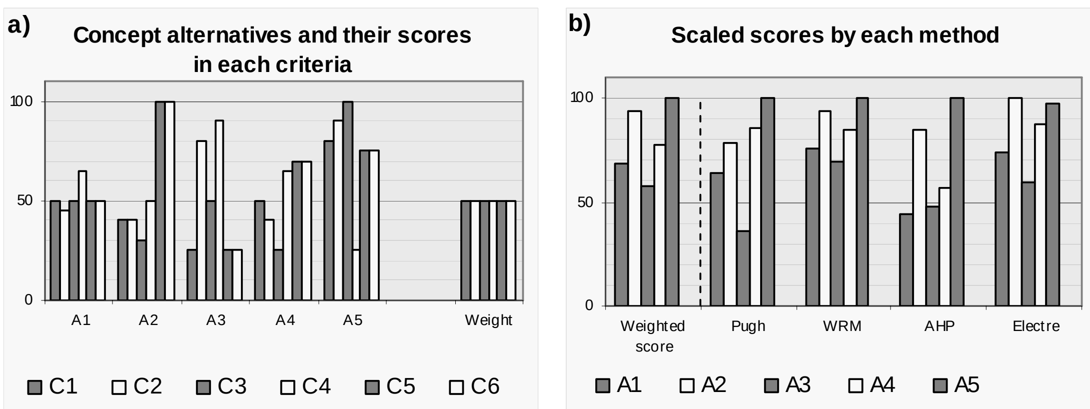

Figure 9. Concept evaluation of "Normal distribution"-case
a) Source value distribution, and b) Evaluation scores

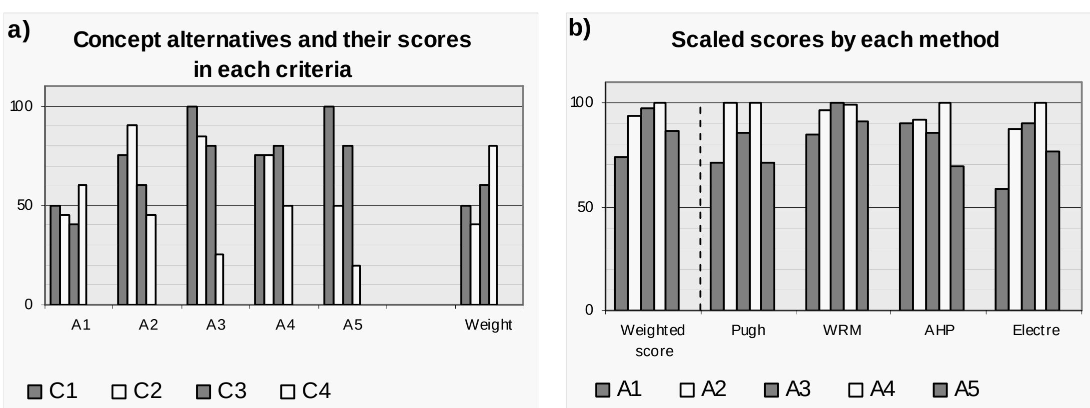

Figure 10. Concept evaluation of "LCD-monitors"-case
a) Source value distribution, and b) Evaluation scores

Results of the comparative study

In some of the cases, the comparison result were clear and all of the methods indicated the same ranking order for the concepts. However, there were also cases where the results varied between methods just as Triantaphyllou [6] has suggested. By examining the alternatives in these cases, we made four notifications.

1) When the full rating scale of AHP was used, the AHP showed the strongest differences between concept alternatives and this can be seen in all of the cases. Even if the weighted average scores of compared alternatives are very close to each other, AHP shows clear differences in scores compared to rest of the methods.
2) Respectively the Electre method showed the smallest relative differences in scores in cases examined in this study.
3) Despite the different calculation method, Electre and WRM produced similar results in cases used in this study. Even more similarities in score distribution can be found from the results of AHP and WRM. The strongest differences in results are between AHP and Electre.
4) In general, the differences in results between the compared methods were relatively small, despite different calculation procedures.

Kamal and Otto&Wood cases are very good examples to study the differences in scores produced by different methods. The Kamal case is very interesting because it shows clear differences between different methods. All except one method indicates alternative three as the most promising one. The distinguishable method is AHP, which clearly indicates concept number four as the best alternative. The reason behind this difference is AHP’s calculation method, which amplifies differences between the alternatives. In the Kamal-case, the criteria weight factors differ greatly from each other and AHP magnifies this. More differences between methods appear in the Otto&Wood case, where the first ranking alternative is quite clear but the second ranking alternative depends on the method used.

4. CONCLUSIONS

In this paper, we have studied the application of multi-criteria decision making methods for concept selection in a comparative study. In the study four concept selection methods were used and compared. These methods were Pugh’s method, Weighted Rating Method, Analytical Hierarchy Process, and Electre III –method. We identified that in certain cases the choice of an MCDM method can have a significant affect on the result.

According to the results of this study, we make five suggestions for users of these methods and for the analysis of the results.

1) Comparison of concept alternatives should be started with the more simple methods e.g. Pugh or weighted rating methods. If there is a reason to question the result, then use of more detailed methods, such as AHP, can be justified.
2) During the definition of weight factors and the evaluation of concept alternatives, one should consider using another method to verify the results, if the weight factors are strongly uneven or if large differences in performance characteristics exist.
3) When the result is very clear with one method, the result most likely will be the same with other methods.
4) When the results show only small differences and concept alternatives can’t be studied and developed further, another method should be used in order to get a wider perspective into results. One can enhance the resolution of ratings, or change the comparison mode from reference value comparison to pair-wise comparison (or vice versa).
5) In addition to performing a comparison, one should always look into the actual comparison data. Check possible best alternative in detail – one criterion at a time – before making the final decision.

We believe that an awareness and understanding of the fundamental differences in the MCDM methods is important for both the designers and the project leaders who are participating in concept selection procedures. This study can help designers to select the most suitable method for their concept selection case, and also improves the analysis of the results given by these methods. The presented list of five suggestions makes the selection of an appropriate method systematic and clear. It also provides certain guidelines for the designers to interpret the comparison results. In further studies, we will concentrate on user perception, and how user interprets source values and performance characteristics to different kind of scales. We also believe that much understanding in this field can be generated and this understanding can help to develop concept selection procedures even further.

REFERENCES

[1] Ulrich K.T., Eppinger S.D., 2000, Product Design and Development, McGraw-Hill, New York, NY, 2nd ed.

[2] Pugh S., 1996, Creative Innovative Products Using Total Design, Addison-Wesley, p.544.
[3] Pahl, G., and Beitz, W., 1996, Engineering Design—A Systematic Approach, 2nd Ed., Springer, London.
[4] Birkhofer H., Jänsch J., Kloberdanz H., 2005, An extensive and detailed view of the application of design methods and methodology in industry, International Conference on Engineering Design, Melbourne, Australia, August 15-18, 2005
[5] Salonen M. and Perttula M. Utilization of concept selection methods – a survey of Finnish industry, In Proceedings of ASME Design Engineering Technical Conferences, Long Beach, CA, USA, September 24-28, 2005.
[6] Triantaphyllou, E., 1995, Using the Analytic Hierarchy Process for Decision Making in Engineering Applications: Some Challenges, International Journal of Industrial Engineering: Applications and Practice, Vol. 2, No. 1, pp. 35-44
[7] Pugh S., 1981, Concept Selection: A Method That Works, Proceedings of the 1981 International Conference on Engineering Design, Rome
[8] Saaty, T.L., 1980, The Analytical Hierarchy Process, McGraw-Hill, New York.
[9] Roy, B., 1991, The Outranking Approach and the Foundation of ELECTRE Methods, Theory and Decision, Vol. 31, pp 49-73
[10] Takai S., Ishii, K., 2004, Modifying Pugh's Design Concept Evaluation Methods, In Proceedings of ASME Design Engineering Technical Conferences, Salt Lake City, Utah, USA, September 28 – October 2, 2004.
[11] B. Roy, D. Bouyssou, 1993, Aide multicritére à la décision: Méthodes et cas, Economica, Paris.
[12] Yeo S.H., Mak M.W., Balon A.P., 2004, Analysis of Decision-making Methodologies for Desirability Score of Conceptual Design, Journal of Engineering Design, Vol. 15, pp. 195-208
[13] Hauser J., Clausing D., 1988, The House of Quality, Harvard Business Review, May-June, pp. 67-73
[14] Brans, J.P., 1982, 'L'ingéniérie de la décision. Elaboration d'instruments d'aide à la décision. Méthode PROMETHEE', Université Laval, Québec, Canada.
[15] Buchanan J, et al., 1999. Project ranking using Electre III. Research Report Series 1999-01, Department of Management Systems.
[16] Kamal M., et al. 2001, Application of the AHP in project management, International Journal of Project Management, Vol. 19, pp 19-27.
[17] Otto, K.N., Wood, K.L., 1995, Estimating Errors in Concept Selection, Design Engineering Technical Conferences, vol. 2, 397-411, 1995.

Contact: Sauli Honkala
Helsinki University of Technology
Machine Design
Otakaari 4, Espoo
PO Box 4100, FI-02015 TKK
Finland
Tel: +358 9 451 3570
Fax: +358 9 451 3549
E-mail: sauli.honkala@tkk.fi
URL: http://www.machina.hut.fi/pdr/