Avoiding MCDA Evaluation Pitfalls

Jonathan Barzilai

Dalhousie University
&
Scientific Metrics

1 Introduction: The Issues

One is not required to be a mechanical engineer to drive a car and, considering the advanced state of mechanical engineering, most people limit their interest in what lies "under the hood" to finding a competent mechanic. Users of evaluation, risk, and decision analysis tools that are based on classical decision theory should be aware that, as is demonstrated below, classical MCDA (multi-criteria decision analysis) has not reached the advanced state of mechanical engineering. Since evaluation and decision tools that are based on flawed mathematical foundations produce meaningless numbers, the purpose of this paper is to give a sample of typical errors and direct the reader to (i) practical tools that are based on sound mathematical foundations and (ii) to these mathematical foundations.

Typically, even the simplest multi-criteria evaluation techniques involve numbers and operations such as addition and multiplication. Also typically, it is not recognized that the numbers on which the operations of addition and multiplication are performed represent preference scales and that these are mathematical operations albeit elementary ones. Although the construction of preference scales and the applicability of mathematical operations to preference scale values are of great theoretical and practical importance, the problem of applicability of these operations has been ignored in the literature following the publication of von Neumann and Morgenstern's *Theory of Games and Economic Behavior* [21] and the conditions under which these operations are applicable have not been identified until recently (see Barzilai [5, 6 and 7]).

2 A Sample of Typical Problems

2.1 The Applicability of Addition and Multiplication

Consider the applicability of the operations of addition and multiplication on scale values for a fixed scale, that is, operations that express facts such as "the weight of a given object equals the sum of the weights of two other objects" ($m(a) = m(b) + m(c)$) and "the weight of a given object is two and a half times the weight of another one"

$(m(a) = 2.5 \times m(b))$. It may be surprising to learn that these operations are not applicable to any scales in the classical literature, but the correct model for preference scales, which are the scales of interest in evaluation, risk, and decision analysis, is that of a straight line and none of the scales in the classical literature is constructed in accordance with the algebraic and geometric structure of the straight line. In fact, the conditions for applicability of addition and multiplication have not been identified in the classical literature and the issue of applicability of mathematical operations cannot be found in the literature following the publication of von Neumann and Morgenstern's book [21].

A technical note: It is important to emphasize the distinction between the application of the operations of addition and multiplication on scale values for a fixed scale, for example $m(a) = m(b) + m(c)$, as opposed to what *appear to be the same* operations when they are applied to an entire scale whereby an equivalent scale is produced by what amounts to a change of the zero point or unit, for example $t = p + q \times s$ where s and t are two scales and p, q are numbers. For details see Barzilai [5].

2.2 On a Scale of 1-10, How Far is Lisbon from Amsterdam?

This is a meaningless question and numbers that are given in answer to similar questions are meaningless as well. Although no mathematical operations are applicable to such numbers, there seems to be nothing in the classical measurement, decision, or evaluation literature to tell marketing experts that there is no basis for the statistical operations which they routinely carry out on numbers received in response to questionnaires that contain such questions. For further details see Barzilai [4].

2.3 Measurement Without Units

Measurement without units produces scales to which addition and multiplication are not applicable (see Barzilai [7, §8]). Yet the term *unit* does not appear in Roberts's *Measurement Theory* [20] and there is no formal definition of the term in the literature. Similarly, there is no formal definition of the term *scale* in *Foundations of Measurement* (Krantz *et al.* [15]).

2.4 Group Decision Making

The common view in the classical literature that group decision making cannot be modelled mathematically is an error that is based on a misinterpretation of the implications of Arrow's Impossibility Theorem [1] (cf. Barzilai [6]). Another approach to group decision making, game theory, cannot serve as a foundation for group decision making as well – see Barzilai [6 and 3].

2.5 Utility Theory

Although utility theory has been the subject of much controversy since its early days, the main flaws in the foundations of this theory have been brought to light only recently. Among other things, the construction phase of utility theory contains a self-contradiction (see Barzilai [6] for details).

2.6 The Analytic Hierarchy Process

More than thirty years after the publication of Miller's work [16, 17, 18], there is still no acknowledgement in the Analytic Hierarchy Process (AHP) literature (or elsewhere) of his contribution to decision theory in general and the AHP in particular. Some of Miller's ideas are valuable while others are mathematically incorrect but almost all of the additions to his original methodology are in error. Many AHP errors are reviewed in Barzilai [8–11] (see also the references there). Not surprisingly, these errors have been mis-identified in the literature and some of these errors appear in decision theory. For example, Kirkwood [14, p. 53] relies on Dyer and Sarin [12] which repeats the common error that the coefficients of a linear value function correspond to relative importance [12, p. 820]. Furthermore, "difference measurement" which is the topic of Dyer and Sarin is not the correct model of preference measurement. As is the case for other preference scales, there is no foundation for the use of the operations of addition and multiplication in the construction of AHP's preference scales (in this case these operations are used to compute the AHP's eigenvector "priorities").

2.7 Pairwise Comparisons and Preference Ratios

Pairwise comparisons (i.e. comparing two alternatives at a time) and ratios of alternatives cannot be used in the construction of preference scales to which the operations of addition and multiplication are applicable. (Until recently, the use of pairwise comparisons and preference ratios was not related in the literature to the applicability of addition and multiplication.) For details see Barzilai [4].

3 Preference Function Modelling

Classical evaluation theories, including utility theory, cannot serve as the mathematical foundation of decision theory, game theory, economics, or other scientific disciplines since they do not enable the operations of algebra and calculus which are needed and widely used in the physical and social sciences and in statistics. A new theory for preference measurement that enables these operations has been developed in Barzilai [6 and 7]. Based on this theory, a practical methodology for constructing proper preference scales, Preference Function Modelling (PFM), and a software tool that implements it, Tetra, have been developed. Tetra requires only simple and intuitive operations and is a powerful tool for group evaluation and decision making. For future developments of the theory, methodology, and software tools consult arxiv.org and www.scientificmetrics.com.

A technical note: In geometrical terms, proper preference scales reflect the objects under measurement to points on a straight line – see Artzy [2], Postnikov [19] and Barzilai [6].

4 Summary

Classical decision theory (e.g. Keeney and Raiffa [13]) and measurement theory (e.g. Krantz *et al.* [15]) are founded on errors that go back to early utility theory and which have been propagated throughout the literature and have led to a proliferation of methodologies and software tools that are based on flawed mathematical foundations and produce meaningless numbers. In addition, the common notion in classical decision theory that group decision making cannot be modelled mathematically is incorrect and is based on results that apply to ordinal systems only.

References

[1] K.J. Arrow, *Social Choice and individual Values,* Wiley, 1951.

[2] R. Artzy, *Linear Geometry,* Addison-Wesley, 1965.

[3] J. Barzilai, "A Game Theory Pitfall – "The Value of a Game"," *CORS-SCRO Bulletin*, p. 44, Vol. 41 No. 3, July 2007.

[4] J. Barzilai, "Pairwise Comparisons," in N.J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*, Sage Publications, Vol. II, pp. 726–727, 2007.

[5] J. Barzilai, "On the Mathematical Modelling of Measurement," http://arxiv.org/, math.GM/0609555.

[6] J. Barzilai, "Preference Modeling in Engineering Design," in *Decision Making in Engineering Design,* K.E. Lewis, W. Chen and L.C. Schmidt (Eds.), ASME Press, pp. 43–47, 2006.

[7] J. Barzilai, "Measurement and Preference Function Modelling," *International Transactions in Operational Research,* Vol. 12, pp. 173–183, 2005.

[8] J. Barzilai, "Notes on the Analytic Hierarchy Process," *Proceedings of the NSF Design and Manufacturing Research Conference,* Tampa, Florida, pp. 1–6, 2001.

[9] J. Barzilai, "On the Decomposition of Value Functions," *Operations Research Letters,* Vol. 22, Nos. 4–5, pp. 159–170, 1998.

[10] J. Barzilai, "Understanding Hierarchical Processes," *Proceedings of the 19th Annual Meeting of the American Society for Engineering Management,* pp. 1–6, 1998.

[11] J. Barzilai, "Deriving Weights from Pairwise Comparison Matrices," *Journal of the Operational Research Society,* Vol. 48, No. 12, pp. 1226–1232, 1997.

[12] J.S. Dyer and R.K. Sarin, "Measurable Multiattribute Value Functions," *Operations Research,* Vol. 27, No. 4, pp. 810–822, 1979.

[13] R.L. Keeney and H. Raiffa, *Decisions with Multiple Objectives,* Wiley, 1976.

[14] C.W. Kirkwood, *Strategic Decision Making,* Duxbury, 1996.

[15] D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky, *Foundations of Measurement,* Vol. 1, Academic Press, 1971.

[16] J.R. Miller, *The Assessment of Worth: A Systematic Procedure and Its Experimental Validation,* doctoral dissertation, M.I.T., 1966.

© J. Barzilai 2007

[17] J.R. Miller, *Assessing Alternate Transportation Systems,* Memorandum RM-5865-DOT, The RAND Corporation, 1969.

[18] J.R. Miller, *Professional Decision-Making,* Praeger, 1970.

[19] M. Postnikov, *Analytic Geometry* (Lectures in Geometry, Semester I), Mir Publishers, 1982.

[20] F.S. Roberts, *Measurement Theory,* Addison-Wesley, 1979.

[21] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior,* Princeton University Press, 1944.