Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Reply to “Remarks on the Analytic Hierarchy Process” by J. S. Dyer

Patrick T. Harker, Luis G. Vargas,

MANAGEMENT SCIENCE
Vol 36, No. 3, March 1990
Printed in U S A.

REPLY TO "REMARKS ON THE ANALYTIC HIERARCHY PROCESS" BY J. S. DYER*

PATRICK T. HARKER AND LUIS G. VARGAS
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6366
Joseph M. Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

The paper by J. S. Dyer (1990, this issue) presents two arguments against the use of the Analytic Hierarchy Process (AHP): the axioms are "flawed" and the rankings which the AHP produces are "arbitrary". In particular, he takes offense of our earlier claim (1987) that much of the criticisms of the AHP are based on a misunderstanding of the theoretical foundations of the AHP. The arguments raised by Dyer are not new (Dyer and Wendell 1985) and were, in fact, the reason for our earlier paper in *Management Science* (Harker and Vargas 1987). In this note, we would like to respond to Dyer's claim that the AHP is "flawed" and to argue that our initial claim is true: this criticism arises out of a lack of understanding of the theory underlying the AHP. However, we shall not respond on a point-by-point basis since this has been done by Saaty (1990, this issue). Thus, we will present a brief discourse on the flaws in Dyer's argument.

Before attacking the axiomatic basis of the AHP and the "flawed" nature of the ranking procedure in the AHP, Dyer claims that the questions asked in the AHP are ambiguous; he writes:

> Suppose a thoughtful person hears the question, "How much better is A_i than A_j on a criterion?" His appropriate response would be, "Relative to what?" This latter question expresses intuitively the need for the definition of the reference point.

First, the questions used in the AHP are not as Dyer describes. Second, the definition of the criterion must always involve a point of reference. Axiom 2 (ρ-homogeneity) implies that for questions to be meaningful, the paired comparisons must be performed on a homogeneous scale. Thus, for example, if we are comparing car A and car B with respect to cost, the appropriate question is: with respect to cost, which of the two cars (A or B) is preferred, and by how much? Of course, we are in agreement with Dyer that if one cannot afford either car, then the response would be meaningless and, thus, the set is not ρ-homogeneous. The very definition of the cost criterion has defined an implicit reference set. Properly considered, the AHP provides a novel means of defining reference points. One would first compare all alternatives to the first (alternative one being a reference point), then compare all to the second alternative, and so on. Thus, the AHP does not take a fixed reference point but, rather, treats all alternatives as reference points in order to minimize any bias which may be introduced through the selection of a single focus for the comparisons.

Dyer also raises a minor objection to the 1–9 scale prior to his main discourse. First, as mentioned in Harker and Vargas (1987), this scale is chosen due to its empirical properties; the theory only requires a bounded ratio scale and that the alternatives are homogeneous with respect to this scale. Dyer presents an example where a decision maker prefers A three times more than B and B five times more than C, which would imply that A be 15 times more preferred than C. However, with a scale bounded by 9, this consistent judgement is not permitted. First, one should point out that alternatives

* Accepted by Robert L. Winkler, former Departmental Editor; received April 7, 1989.

0025-1909/90/3603/0269$01.25

A and C are not homogeneous according to the 1–9 scale; a different scale is necessary if 15 is to be permitted. Second, even if one were forced to compare these alternatives within a 1–9 scale, not much difference arises:

Consistent Matrix	Weight	Approximate Matrix	Weight
1 3 15	0.714	1 5 9	0.716
1 5	0.238	1 8	0.235
1	0.048	1	0.049

Therefore, the use of the 1–9 scale does not affect the theory of the AHP, empirical evidence suggests that it appears to be an appropriate scale to capture decision maker's preferences, and as the above example suggests, the AHP is fairly robust even if "errors" are made.

1. Are the Axioms of the AHP Valid?

Dyer criticizes the axioms of the AHP as not following from basic axioms but, rather, from "primitive notions" which includes the assumption of a fundamental scale. We must first point out that Dyer fails to recognize that the axioms of utility theory (UT) are based on more than the use of a binary relation $\succ$. In addition to this concept, one must further define this relation on a set of lotteries which in turn assumes the existence of a scale—the probability scale. Thus, the existence of a von Neumann-Morgenstern utility function is a consequence of the existence of probabilities. In the case of ordinal utility functions, one can define a fairly weak fundamental scale. Nonordinal scales, however, require further refinement of the fundamental scale. In UT, this results in an interval scale; in the AHP, one has a ratio scale. Why an interval scale? The choice of the scale is a question of belief, not of "truth" as Dyer would have one believe. Also, is the notion of probability so intuitive to humans that it must be taken as the basis of all decision theories? Allais' paradox and other such results from experimental decision analysis cast doubt on this claim.

Furthermore, Dyer claims that UT follows from axioms and not on the assumed existence of a scale of measure (a primitive notion in Dyer's language). Both UT and the AHP have primitive notions since it is impossible to derive strengths of preferences from simple axioms of order. Thus, the existence of paired comparisons in the AHP play the same role as lotteries do in UT. The basic difference is that in UT, transitivity is a necessary condition while in the AHP, it is not. Both methods must assume the existence of a scale (ratio or interval) and thus, it is simply not clear why one would want to follow Dyer's advice to base all AHP results on the axioms of UT.

In simple words, Axiom 1 of the AHP states that humans can make comparisons and in this process, we can use ratio scales. Note that this assumption does not require the use of negative numbers which must be defined when interval scales are used. Are interval scales the only way to think about the world as Dyer would have one believe? The choice of scale is much more complex than is depicted by Dyer; we believe that ratio scales provide a very useful way to model a wide variety of situations and thus, one should not simply assume that interval scales are the only method to measure cognition.

Axiom 2 deals with our limited cognitive capabilities. It is unrealistic to believe that humans can effectively employ unbounded perceptual scales; this axiom states that a limit exists on our ability to compare widely differing alternatives or criteria.

The remaining axioms are designed to deal with multiple criteria. Why are these axioms any less appealing than those involving lotteries? We sincerely doubt that the majority of people would find lotteries more intuitive than the concepts described by the axioms of the AHP.

2. Rank Reversals

The notion that the AHP violates independence of irrelevant alternatives (rank reversal) is the major place where Dyer believes the AHP to be "flawed". This topic was considered at length in our earlier paper (Harker and Vargas 1987) and hence, we will focus our attention on the situations in which the principle of hierarchic composition is applicable.

The only reason Dyer gives for the AHP yielding "arbitrary rankings" is that the AHP does not possess an independence axiom; this is not the case. Axiom 3 states very clearly what independence means in the context of the AHP. Dyer proposes the use of the concept of difference independence to rectify the problem he perceives in the AHP. *All* concepts of independence in both UT and the AHP are subjective; they are based on a simple idea: do our preferences change when some of the criteria change? In the example given by Dyer, the appropriate question is: given a pair of criteria C_i, C_j, if C_i is preferred to C_j when considering alternative A_k, can we conclude that C_i is preferred to C_j for all alternatives? In the AHP, preference is not just an ordinal scale. Thus, the strength of preference of C_i over C_j must be preserved from alternative to alternative. UT never discusses the question of the independence of the criteria from the alternatives. However, the question is as applicable in that theory as it is in the AHP; the AHP has simply provided a formal mechanism for its investigation. Since criteria are experienced through alternatives, it seems difficult to separate the criteria from the alternatives. In the example provided by Dyer, we observe that the importance of the criteria vary from alternative to alternative. Axiom 3 of the AHP states that for hierarchies (the only structure for which the principle of hierarchic composition is valid), a level does not depend on the level above it in the sense mentioned above. In addition, the elements within a level must be independent among themselves. It is clear that Dyer's example does not satisfy Axiom 3.

The solution to the example presented by Dyer is to use a system with feedback (Harker and Vargas 1987). Dyer dislikes this solution:

> The problem of the large number of required ratio comparisons is compounded by the difficulty of responding to the question requiring that the criteria be compared with respect to the alternatives.

However, his critique fails to formulate the correct question to compare the criteria to the alternatives:

> . . . the decision maker would be required to respond to questions such as this, "Does the Mercedes perform better on cost or on appearance, and by how much?"

This question is simply a misunderstanding on the part of Dyer. The proper question would be: given an alternative (e.g., the Mercedes), which criterion (e.g., appearance or cost) is more important in the overall choice of the best car? A decision maker may respond with the same relative importance of each criterion no matter which alternative is selected; in this case, hierarchic composition can be used. If the answers vary with the chosen alternative (the reference point), then the supermatrix technique described in Harker and Vargas (1987) and Saaty (1990, this issue) must be employed; the problem is not a hierarchy! While Dyer may doubt that such questions can be answered, empirical applications (Hämäläinen and Seppäläinen 1986) of this technique provide evidence that the questions are meaningful to decision makers. Furthermore, this technique provides a test for whether or not a hierarchy can be used to model the given decision problem. Thus, Dyer's claim that independence cannot be tested in the AHP is false.

3. Comments on Dyer's Proposed Method

After discussing the rank reversal issue, Dyer states:

> A remarkable observation regarding this controversy about the phenomenon of rank reversal in the AHP is that a simple solution does exist. . . . This solution is appropriate as long as the criteria used in the application of the AHP satisfy the property of difference independence . . .

The method proposed by Dyer does not work because it is possible to show that it does not truly capture the ranking of alternatives when scales with a standard unit are available to measure alternatives according to criteria. What happens when no scales exist (and hence the AHP relative scales are used) for the criteria or when there are subcriteria?

Here is a simple example where Dyer's proposed method does not work. We have three alternatives measured with the same scale with respect to two criteria:

$$\begin{array}{ccc} & & \underline{\text{Criteria}} \\ & & \begin{array}{cc} C_1 & C_2 \end{array} \\ \text{Alternatives} & \begin{array}{c} A \\ B \\ C \end{array} & \left[\begin{array}{cc} 80 & 1 \\ 79 & 7 \\ 75 & 8 \end{array}\right] \begin{array}{c} 81 \\ 86 \\ 83 \end{array} \\ & & \begin{array}{cc} 234 & 16 \end{array} \end{array}$$

The criteria weights are the relative totals for each of the two columns:

$$\underline{\text{Criteria Weights}}$$
$$\left[\begin{array}{c} 234/250 = 0.936 \\ 16/250 = 0.064 \end{array}\right]$$

Thus the relative weights of the alternatives are given by:

$$\begin{array}{cc} \underline{\text{Relative Weights}} & \underline{\text{Rank}} \\ \begin{array}{c} A \\ B \\ C \end{array} \left[\begin{array}{c} 81/250 = 0.324 \\ 86/250 = 0.344 \\ 83/250 = 0.332 \end{array}\right] & \begin{array}{c} 3 \\ 1 \\ 2 \end{array} \end{array}$$

According to Dyer the utilities of the alternatives under each criterion are obtained by subtracting from each alternative value under a criterion the lower bound of the scale and dividing by the difference of the upper and lower bounds of the scale, i.e., $(x - 1)/(80 - 1)$. When applied to the above we get on using instead of the x entry above:

$$\begin{array}{cccccc} & & \underline{\text{Criteria}} & \underline{\text{Criteria Weights}} & \underline{\text{Utilities}} & \underline{\text{Rank}} \\ & & \begin{array}{cc} C_1 & C_2 \end{array} & & & \\ \text{Alternatives} & \begin{array}{c} A \\ B \\ C \end{array} & \left[\begin{array}{cc} 1 & 0 \\ .9873 & .0759 \\ .9367 & .0886 \end{array}\right] & \left[\begin{array}{c} 234/250 = 0.936 \\ 16/250 = 0.064 \end{array}\right] & = \begin{array}{c} A \\ B \\ C \end{array} \left[\begin{array}{c} 0.936 \\ 0.929 \\ 0.882 \end{array}\right] & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \end{array}$$

Note that the utilities do not yield the same rank as the relative weights which give the true ranks of the alternatives.

It is pointless to preserve rank with respect to a method that yields false results.

Even if Dyer were to remedy this situation, he says nothing about what one should do when there are no scales with a unit and we must generate relative weights for the alternatives from the start. In that case rank reversal could occur because the measurement of the alternatives depends on those considered.

4. Summary and Conclusions

Let us summarize the central element of the discussion concerned with rank reversal. If we have a single criterion such as redness, and apple A is redder than apple B, then one would object to a new apple C causing apple B to become redder that A according to some theory. Indeed this cannot happen in the AHP when the judgments are consistent. However, it can in principle in UT since this theory requires the subjective notion of an irrelevant alternative for remedy. If an alternative is irrelevant it cannot cause rank reversal. Note that in the AHP every alternative that can be compared with other alternatives can be called relevant in that terminology. When Utility Theory (a single-criterion theory) was developed, it was noticed that rank reversal could not be guaranteed. The same problem was inherited by the more general Multi-Attribute Utility Theory. This concern has led to bizarre attempts by advocates of Utility Theory to preserve rank at all costs and without rigorous justification. If we have several criteria, then the traditional assumption is that rank reversal should be allowed only if new criteria are introduced or the weights of the old criteria are changed when a new alternative is added or an old one deleted. The AHP suggests that one must decide whether to preserve rank due to acquired standards of what "should or should not" be by using the ranking procedure, or allow rank to take its natural course by making paired comparisons. The reason why rank can reverse in the AHP with relative measurement is clear. It is because the alternatives depend on what alternatives are considered, hence, adding or deleting alternatives can lead to change in the final rank. Thus, contrary to Utility Theory, no mathematically unjustified rank reversal occurs in the AHP. Utility theorists should direct their energy to preserving rank in their theory in a mathematically justifiable way rather than banning rank reversals from the domain of what constitutes rational behavior.

References

DYER, J. S., "Remarks on the Analytic Hierarchy Process," *Management Sci*., 36, 3 (1990), 249–258.

——— AND R. E. WENDELL, "A Critique of the Analytic Hierarchy Process," Working Paper 84/85-4-24, Department of Management, The University of Texas at Austin, 1985.

HÄMÄLÄINEN, R. P. AND T. O. SEPPÄLÄINEN, "The Analytic Network Process in Energy Policy Planning," *Socio-Economic Planning Sci*., 20 (1986), 388–405.

HARKER, P. T. AND L. G. VARGAS, "The Theory of Ratio Scale Estimation: Saaty's Analytic Hierarchy Process," *Management Sci*., 33 (1987), 1383–1403.

SAATY, T. L., "An Exposition of the AHP in Reply to the Paper: Remarks on the Analytic Hierarchy Process," *Management Sci*., 36, 3 (1990), 259–268.