Article:

Ishizaka, A and Siraj, S orcid.org/0000-0002-7962-9930 (2020) Interactive consistency correction in the analytic hierarchy process to preserve ranks. Decisions in Economics and Finance, 43 (2). pp. 443-464. ISSN 1593-8883

https://doi.org/10.1007/s10203-020-00309-4

Interactive consistency correction in the analytic hierarchy process to preserve ranks

Alessio Ishizaka[1] , Sajid Siraj[2]

[1]NEOMA Business School, 1 rue du Maréchal Juin - BP 215 - 76130 Mont-Saint-Aignan Cedex – France, alessio.ishizaka@neoma-bs.fr
[2]Leeds University Business School, The University of Leeds, LS2 9JT, United Kingdom, s.siraj@leeds.ac.uk

Abstract: The analytic hierarchy process is a widely used multi-criteria decision-making method that involves the construction of pairwise comparison matrices. To infer a decision, a consistent or near consistent matrix is desired, and therefore several methods have been developed to control or improve the overall consistency of the matrix. However, controlling the overall consistency does not necessarily prevent having strong local inconsistencies. Local inconsistencies are local distortions which can lead to rank reversal when a new alternative is added or deleted. To address this problem, we propose an algorithm for controlling the inconsistency during the construction of the pairwise comparison matrix. The proposed algorithm assists decision makers whilst entering their judgments and does not allow strong local inconsistencies. This algorithm is based on the transitivity rule and has been verified through statistical simulations. Appropriate thresholds of acceptable evaluations have been inferred from these simulations. We demonstrate that the proposed algorithm is a helpful decision aid to decision makers when entering pairwise comparison judgments.

Keywords: Decision making; Multi-criteria; AHP; Rank reversal; Consistency ratio

1. Introduction

The analytic hierarchy process (AHP) is a structured method to help people in making complex yet important decisions. It provides a comprehensive and rational framework to structure a problem, represent and quantify its elements, relate those elements to the overall goals, and finally evaluate and prioritise the available alternatives (Ishizaka, Labib 2011a; Saaty 1980). The AHP method has been extensively used for making strategic decisions in the presence of conflicting criteria (Sipahi, Timor 2010; Subramanian, Ramanathan 2012). It has also been used for performance measurement and efficiency analysis, either as a standalone method (Ishizaka, López 2019; Ishizaka, Labib 2011b) or in combination with data envelopment analysis (DEA) (Tone 2017; Veni et al. 2012; Keskin, Köksal 2019). Besides the fact that it has been widely applied, it has been widely criticised in the academic community for the rank reversal problem. A rank reversal occurs whenever the relative ranking of two alternatives in the global ranking is reversed when a new alternative is added or removed. In fact, the rank reversal can be produced by two distinct and independent causes: the normalisation of the weights and strong locally inconsistent pairwise comparisons. The first cause has been largely debated (Maleki, Zahir 2012). One side has criticised the rank reversal (Schoner et al. 1992; Dyer 1990b, a; Holder 1990, 1991; Schoner, Wedley 1989; Triantaphyllou 2001) because, in a normative approach, it violates the axiom of independence: once it has been decided that A is better than B, this statement must remain true independent of the number of other alternatives that will be considered. Another part has legitimised it with the illustration of several real-life examples, where the introduction of a new alternative may lead to a rank reversal (Harker, Vargas 1987, 1990; Pérez 1995; Saaty, Takizawa 1986; Saaty 1987, 1990; Saaty 1991, 2006; Saaty, Sagir 2008). As discussed in (Millet, Saaty 2000; Saaty, Sagir 2008), two different modes can be used to normalise the scores in AHP: the distributive mode and the ideal mode. The choice of which mode to use in a given situation depends on whether rank reversal is acceptable or not. The distributive mode, that does not preserve ranks, uses a normalisation by dividing the score of each alternative by the sum of all local priorities. The sum of the new local priorities will be one. The ideal mode that preserves ranks in rating alternatives, under the conditions that the best alternative is not deleted or a better alternative is not added, uses a normalisation by dividing the score of each alternative only by the score of the best alternative under each criterion.

In this paper, we will first review the main results of this debate in Section 2. In Section 3, we will study the second cause of rank reversal that has not been yet investigated in the literature. We will present an algorithm which controls inconsistent pairwise comparisons and therefore avoids rank reversals due to strong local inconsistencies. These improvements to the method will render it more reliable as a performance analysis tool. Section 4 verifies the algorithm with a statistical simulation, and finally Section 5 concludes the paper.

2. Rank reversal due to normalisation of weights

In multi-criteria problems, a normalisation is needed when criteria are measured on different scales, for example if one scale is in centimetre and the second is in metres, then a normalisation is needed to rescale 100 cm to 1 metre. The drawback is that a rank reversal due to the normalisation of the weights is a common problem of all multi-criteria decision methods based on the weighted sum model. For example, suppose that we have three criteria $C_i,\ i = 1{,}2{,}3$ with equal weights (that is, $w_i = \frac{1}{3}$) and we have three alternatives $A_j,\ i = 1{,}2{,}3$. We further assume that the scores have already been received for alternatives A_j with respect to each criterion C_i. Let us assume that all these scores are on an absolute scale of 1 to 10, as given in Table 1.

	C_1	C_2	C_3
A_1	1	9	8
A_2	9	1	9
A_3	1	1	1

Table 1: Scores of alternatives with respect to each criterion

Each of the columns in Table 1 can be normalised using the following standard formula:

$$\bar{s}_{ij} = \frac{s_{ij}}{\sum_j s_{ij}}$$

If we normalise the scores using this approach and aggregate them using the weighted sum formula (i.e. $v_j = \sum_i w_i \bar{s}_{ij}$) , then we obtain the overall scores (v_j) as given below:

v_1: 1/11 · 1/3 + 9/11 · 1/3 + 8/18 · 1/3 = 0.45

v_2: 9/11 · 1/3 + 1/11 · 1/3 + 9/18 · 1/3 = 0.47

v_3: 1/11 · 1/3 + 1/11 · 1/3 + 1/18 · 1/3 = 0.08

which give the final ranking: $v_2 > v_1 > v_3$.

We now suppose that a new alternative A_4, which is a copy of the alternative A_2, is added for consideration (Table 2).

	C_1	C_2	C_3
A_1	1	9	8
A_2	9	1	9
A_3	1	1	1
A_4	9	1	9

Table 2: Scores of alternatives along with additional A_4 with respect to each criterion

If we now normalise the scores of each column to one and aggregate them again, we obtain:

v_1: $1/20 \cdot 1/3 + 9/12 \cdot 1/3 + 8/27 \cdot 1/3 = 0.37$

v_2: $9/20 \cdot 1/3 + 1/12 \cdot 1/3 + 9/27 \cdot 1/3 = 0.29$

v_3: $1/20 \cdot 1/3 + 1/12 \cdot 1/3 + 1/27 \cdot 1/3 = 0.06$

v_4: $9/20 \cdot 1/3 + 1/12 \cdot 1/3 + 9/27 \cdot 1/3 = 0.29$

which give the final ranking: $v_1 > v_2 = v_4 > v_3$.

The final ranking has been reversed with the introduction of a new alternative. Belton & Gear (1983) introduced a similar example, where the scores of Tables 1 and 2 were determined using pairwise comparisons (and not directly with an absolute measurement), as we have done. Other examples of rank reversals can also been found for TOPSIS (García-Cascales, Lamata 2012; Wang, Luo 2009) and the Borda-Kendall method (Wang, Luo 2009).

The rank reversal is due to different normalisation outcomes of the values in Tables 1 and 2; therefore, it is not a specific problem of AHP. A different normalisation and a related rank reversal is also observable in the case of the suppression of an alternative (Troutt 1988) or the introduction of a new alternative which is not a copy of an existing one (Dyer 1990b). For the same reason, a rank reversal may occur when a wash criterion is deleted (Finan, Hurley 2002) or an indifferent criterion is added (Perez et al. 2006). However, the exercise of adding or deleting criteria has also been criticised (Saaty,Vargas 2006).

From a theoretical point of view, it is better to have a model whose results are normalisation-independent as it would avoid a rank reversal. Therefore, it was suggested to use a normalisation "by dividing the score of each alternative only by the score of the best alternative under each criterion" (Belton, Gear 1983). Later in the literature, this normalisation is called B-G normalisation or ideal mode. A counterexample was provided in (Saaty, Vargas 1984) to show that the ideal mode is also subject to rank reversal. In this case, they introduce an alternative, which is a copy in only two of three criteria. In the last criterion, the new alternative has the largest value, which implies different normalisation and priority outcomes. Belton and Gear (1985) responded that if a new alternative is introduced, then the weight criteria should also be modified. This contradicts the AHP philosophy of independence of weights and alternatives. Finally, Vargas (1985) gave a simple example where known priorities can only be retrieved with the distributive mode.

Later, another approach was proposed as the linking pin AHP (Schoner et al. 1993) wherein the local priorities of a specific alternative are normalised to unity. The rank is preserved because the normalisation is always the same. However, the final solution depends on which alternative is selected to link across criteria, and a rank reversal can still occur if the specific alternative used for the normalisation is removed. Recently, it has been proposed (Wang, Elhag 2006) to normalise by the sum of all priorities with the exception of the newly added one. This method is, of course, not applicable to a real case where the process of adding/deleting alternatives should be the same as if when starting anew.

To avoid the normalisation problem, a multiplicative aggregation (later called the multiplicative analytic hierarchy process) has been proposed (Lootsma 1993; Barzilai, Lootsma 1997). A rescaling of any variable has no effect on the results because the overall product will be affected by the same factor for each alternative. For example, if all values of a particular criterion are divided by 100, this does not affect their relative ranking as the overall scores remain proportional to what they were previously. However, in a multiplicative aggregation, compromises are preferred over extremes

(Ishizaka et al. 2010). For example, if a person is indifferent to apples or pears, (s)he will prefer a bag A with 5 apples and 4 pears (total score 5•4 = 20) compared to a bag B with 15 apples and 1 pear (total score 15•1 = 15). This seems illogical as the recommended alternative A (9 fruits) has a far lower number of fruits than alternative B (16 fruits) has. In another example, Vargas (1997) demonstrates that the additive aggregation can retrieve the exact priorities, which is not the case of a multiplicative aggregation. Furthermore, Saaty,Vargas (2006) mention that in the case of weights lower than 1, the calculated ranking would not be correct. For example, we consider two criteria with weights 0.3 and 0.7, respectively, and an alternative with a score of 0.5. The weighted scores are $0.5^{0.3} > 0.5^{0.7}$, which is a contradiction because an alternative has a larger weighted score under a lower priority score.

The cause of a rank reversal is the variation of the normalisation when we add or delete an alternative. Therefore, we need a normalisation invariant to the number of alternatives. Such normalisation can only be given by an alternative that will not be deleted, for example a reference alternative, which is called an ideal alternative in the literature. However, sometimes a rank reversal due to the introduction or the deletion of an alternative is desired as it represents a real-world phenomenon. Rank reversals have been observed experimentally many times (a review can be found in (Tversky et al. 1990; Camerer 1995), and several practical examples are given in (Saaty, Sagir 2008). The most famous is borrowed from (Corbin, Marley 1974): "A lady in a small town wishes to buy a hat. She enters the only hat store in town, and finds two hats, *a* and *b*, that she likes equally well, and so might be considered equally likely to buy. However, now suppose that the sales clerk discovers a third hat *b′*, identical to *b*. Then, the lady may well choose hat *a* for sure (rather than risk the possibility of seeing someone wearing a hat just like hers)."

Millet and Saaty (2000) and Forman and Gass (2001) gave some guidance on which normalisation to use.

a) If we are in a closed system (i.e. no alternative will be added or removed), then the distributive mode should be used.

b) If we are in an open system (i.e. alternatives can be added or removed) and we allow our preferences for alternatives to be dependent on other alternatives (i.e. rank reversal phenomenon is accepted), then the distributive mode is indicated.

c) If we are in an open system and you do not want other alternatives to affect the outcome, then the ideal mode is recommended.

Both modes are implemented in Expert Choice and MakeItRational, the leading software packages supporting AHP. The final decision of which normalisation mode to use is left to the decision maker. Therefore, it is important that the decision maker clearly understands the normalisation process and its consequences as it can lead to different rankings.

3. Rank reversal due to inconsistent pairwise comparisons

3.1 Consistency

In AHP, the assessment of preferences of n decision elements (criteria or alternatives) is performed through pairwise comparisons. These pairwise comparisons are recorded in a reciprocal matrix, $M = (a_{ij})$, i, j=1,..., n, where a_{ij} is a pairwise comparison comparing the i^{th} decision element to the j^{th} decision element. If all a_{ij} in the matrix M respect the transitivity rule (1), then M is said to be (fully) consistent.

$$a_{ij} = a_{ik} \cdot a_{kj} \tag{1}$$

where i, j and k range from 1 to n, representing any decision element of the matrix.

Another way to define a (fully) consistent matrix is when all comparisons a_{ij} between alternatives i, j are exactly the ratio of the priorities w_i and w_j of the compared alternatives i, j:

$$a_{ij} = w_i/w_j \tag{2}$$

As priorities make sense only if derived from consistent or near consistent matrices, a consistency check must be applied. For this task, Saaty (1977) proposed the consistency index (CI):

$$\text{CI} = \frac{\lambda_{\max} - n}{n-1}, \tag{3}$$

where n : dimension of the matrix

λ_{max} : maximal eigenvalue

Customarily, if the consistency ratio (CR), the ratio of CI and RCI (the average CI of 500 randomly filled matrices), is less than 10%, then the matrix can be considered as having an acceptable consistency.

$$CR = CI/RCI, \tag{4}$$

where CR : consistency ratio

RCI : random consistency index

Other consistency indices exist as compiled in (Brunelli 2018). They share high similarity in their results (Brunelli et al. 2013).

3.2 Rank reversal

If a matrix is inconsistent, there can be a contradiction between the comparison of two alternatives and the derived ranking. This phenomenon is also visible in PROMOTHEE II (De Keyser, Peeters 1996) and ELECTRE II and III (Wanga, Triantaphyllou 2008; Figueira, Roy 2009), where it is called pairwise rank reversal (Mareschal et al. 2008). The pairwise rank reversal in AHP was also observed in (Bana e Costa,Vansnick 2008), where it was declared that this phenomenon is unacceptable because it violates the condition of order preservation. Surprisingly, the authors missed the cause of the pairwise rank reversal and attribute it to the eigenvalue derivation method instead of the inconsistency in the comparison matrix (Wang et al. 2009).

Let us now analyse the relation between a pairwise rank reversal and a rank reversal. If we consider an inconsistent matrix with a pairwise rank reversal, i.e. there are at least two alternatives a and b such that b is preferred to a, and a has a higher rank than b in the derived ranking of the comparison matrix. If we delete one alternative other than a and b, the final relative position of a and b can be swapped (see Example 1).

Example 1:

Consider the comparison matrix of Table 3 with a consistency ratio of 0.07. This matrix has a pairwise rank reversal as the comparison c_{ab} = 1/2 but in the derived ranking *a* is on a higher rank than *b*. If we delete the alternative *e* (Table 4), we have a rank reversal as *b* is ranked higher than *a*.

	a	*B*	*c*	*d*	*e*	Priority	Rank
a	1	**1/2**	3	5	4	0.341	**1**
b	2	1	2	4	1	0.313	**2**
c	1/3	1/2	1	2	1	0.128	4
d	1/5	1/4	1/2	1	1/2	0.066	5
e	1/4	1	1	2	1	0.152	3

Table 3: Slight inconsistent comparison matrix and the derived priorities, CR = 0.07

	a	*b*	*c*	*d*	Priority	rank
a	1	1/2	3	5	0.344	**2**
b	2	1	2	4	0.421	**1**
c	1/3	1/2	1	2	0.155	3
d	1/5	1/4	1/2	1	0.080	4

Table 4: Comparison matrix without the alternative *e*, CR = 0.05

A consistency ratio (4) below 10% is not a guarantee to avoid rank reversals because it measures an average inconsistency of the matrix and cannot detect local inconsistencies. For example, in the matrix of Table 3, the relation $c_{ab} \cdot c_{be} \neq c_{ae}$ does not satisfy the transitive relation (1) in intensity and even in direction. In fact, $c_{ab} \cdot c_{be} = 1/2 \cdot 1 = 1/2$, which has an opposite direction of $c_{ae} = 4$. We believe that it is improbable that such inconsistency exists in the mind of the decision maker. The transitivity in AHP means that a strict multiplication should be respected. But, we must respect the fact that if we have c_{ab} and c_{be}, then c_{ae} must be a value coherent with them. Intransitivity may occur due to the complexity of the problem; for example, the higher the number of items to be compared, the greater the chance to experience fluctuation in the attention of decision makers. In such cases, the items may differ by an amount which falls below the just noticeable difference, or simply the decision maker becomes a bad judge (Kendall, Smith 1940). Whatever the reason may be, (s)he should at least be warned of this circumspect relation. In the next section, we will discuss how to detect such irregularities and improve the consistency.

3.3 Methods to improve the consistency

Methods were developed to improve the consistency once the full matrix has received all the comparisons (Saaty 2003; Ergu et al. 2011; Ergu et al. 2013). These methods search for the comparison which contributes the most to the inconsistency. They do not necessarily spot the same comparison. The decision maker may change this most inconsistent comparison or consider the next most inconsistent judgment. However (s)he has the freedom not to change the most inconsistent comparison and to consider successive comparisons. The decision maker may change it under the pressure of the software, which may therefore result to an unwanted ranking. Moreover, the decision maker does not receive any explanation of the reasons of the inconsistency.

Heuristic approaches (Cao et al. 2008; Xu, Wei 1999; González-Pachón, Romero 2004) that auto-generate a consistent matrix from an inconsistent matrix were introduced. However, these approaches must be used carefully as the decision maker has no control on the adjustment of comparisons. It has been shown in experiments (Linares 2009; Gaul, Gastes 2012) that subjects prefer the priorities calculated from the inconsistent matrix rather the priorities of the artificially fixed consistent matrix. Finally, interactive methods have been proposed. The simplest algorithm detects ordinally intransitive comparisons that can be overruled by the decision maker in case of disagreement with the suggested changes (Siraj et al. 2012; Kwiesielewicz, van Uden 2004). The next section presents an improvement of this work by interactively detecting cardinally intransitive comparisons.

3.4 Interactive method for consistency control

Unlike the existing method, where the algorithm to improve consistency is applied after acquiring the complete matrix, the interactive method intervenes during the process of constructing the matrix. In this way, if an evaluation is inconsistent with the previously given evaluations, suggestions on how to correct the inconsistency are provided. At the end of the construction process, we will have a consistent or near consistent matrix. The first interactive method has been proposed in (Ishizaka, Lusti 2004). When filling a comparison matrix, we distinguish two types of comparisons: the independent comparisons (right part of the relation (1)) and the transitive comparisons (left part of the relation (1)). The algorithm checking the inconsistency should first distinguish between these two types of comparisons. Independent comparisons do not require any further action. After entering a few entries, some constraints on further entries spontaneously emerge. These transitive comparisons are assessed based on their compatibility with the previously entered comparisons. The order in which the independent comparisons are chosen may matter. The overall outcome may well depend on which $n-1$ independent components are entered. This first algorithm was to impose the first diagonal to contain the independent evaluations. This can be tedious because a decision maker may not be familiar with these comparisons. Therefore, in this paper we develop a model where any comparison that cannot be derived from previously entered comparisons can be defined.

In the previous interactive method (Ishizaka, Lusti 2004), a tolerance error was allowed in the transitivity rule (1). However, the threshold of this error was not defined. In this paper, we also define, with simulations, what could be an acceptable threshold.

In the next sections, we describe the new improved interactive algorithm.

3.4.1 Independent comparisons

In a reciprocal $n \times n$ matrix there are exactly $n-1$ independent comparisons. These $n-1$ independent comparisons are the minimum non-redundant information required to calculate priorities. A comparison is independent if it contains new information which cannot be derived from other comparisons by the transitivity rule (1). This means that a_{ij} gives a judgment between two items, where least one has never been evaluated before. Practically for the user, a comparison must be added in a new column or row, where, with the exception of the principal diagonal, no independent comparison or reciprocal independent comparison has been made (Table 5).

1	2	6	a_{14}
1/2	1	a_{23}	a_{24}

1/6		1	a_{34}
			1

Table 5: In this matrix, the comparison a_{23} is not independent because row 2 and column 3 already contain a comparison. In fact, we can calculate $a_{23} = a_{21} \cdot a_{13}$. The last independent comparison is to be chosen in the last column between a_{14}, a_{24} or a_{34}.

3.4.2 Transitive comparisons

Transitive comparisons can be deduced from the independent comparisons. For example, in Table 6 the transitive comparisons are

$$a_{14} = a_{12} \cdot a_{24} = 2 \cdot 1/4 = 1/2$$

$$a_{23} = a_{21} \cdot a_{13} = 1/2 \cdot 6 = 3$$

$$a_{34} = a_{31} \cdot a_{12} \cdot a_{24} = 1/6 \cdot 2 \cdot 1/4 = 1/12$$

1	2	6	a_{14}
1/2	1	a_{23}	1/4
1/6		1	a_{34}
	4		1

Table 6: All the transitive comparisons a_{14}, a_{23} and a_{34} can be deduced from the independent comparisons.

As (2), we have to find a combination of independent comparisons providing this ratio of weights, like in example 2.

Example 2:

$$a_{ij} = a_{ix} \cdot a_{x.} \cdot \ldots \cdot a_{.z} \cdot a_{zj} = w_i/\cancel{w_x} \cdot \cancel{w_x}/\ldots \cdot \ldots \cdot \ldots/\cancel{w_z} \cdot \cancel{w_z}/w_j = w_i/w_j \qquad (5)$$

Practically, we start with a comparison having the same numerator as the denominator of the researched transitive comparison. Then, we search for another comparison which has the same numerator as the denominator of the previous comparison. The process stops when we have a comparison with a denominator equal to the denominator of the transitive comparison.

A tabu list (this is not related to the tabu seach algorithm) is written in order to avoid a recursive loop. It contains the used comparisons and their reciprocal values (see Example 3). They must not be reused because it would cause a recursive loop. The following algorithm describes how to find the value of a_{ij}. This can be easily programmed in Prolog because it has a backtrack functionality, which easily allows searches for other solutions.

```
1   Empty the tabu list
2   Row[H|T] = row i of the matrix
3   Initialise buffer =1
4   column = 1
5   SearchRow (Row[H|T]) {
6           IF Row[empty] THEN
7                   Backtrack
8           ELSE IF Row[H]≠ 0 AND H is not in the tabu list THEN
9                   buffer = buffer • H
10                  IF column = j THEN
11                          aij = buffer
12                          EXIT
13                  ELSE
14                          Set H and its reciprocal in the tabu list
15                          Row[H|T] = row column of the matrix
16                          column = 1
17                          SearchColumn(Row[H|T]))
18                  END IF
19          ELSE
20                  column ++
21                  SearchRow(Row[T])
22          END IF
23  }
```

<u>Explanations</u>:

Lines	Explanations
1–3	Initialisations

2	All the elements of the row a_{ij} are copied in the list Row[H|T], where H is the head, i.e. first element, and T is the tail, i.e. the other elements; (missing comparisons are set to 0).
3	*buffer* is a temporary variable for the partial result.
4	*column* notes in which column we are searching.
5	SearchRow(list) is a module to search in the row of the matrix.
6–7	Impasse: no solution is found, and we backtrack to find an alternative.
8	We search for the first entered comparison, which is not in the tabu list; if a backtrack is used in 7, a second comparison is searched from here.
10–12	If the column of the newly found comparison is the same as the searched transitive comparison, then we have the result and we exit the algorithm.
14–18	The found comparison is set in the tabu list; the module SearchColumn(list) is launched with, as argument, the row j corresponding to the column j of the found comparison a_{ij}.
20–21	We do not have an independent comparison in this column. We search in the next column.

Example 3:

The following matrix is given with the independent comparisons, and the transitive comparison a_{35} is required (Table 7):

1	$a_{12} = 5$		$a_{14} = 2$	
$a_{21} = 1/5$	1	$a_{23} = 1/2$		$a_{25} = 3$
	$a_{32} = 2$	1		$\boldsymbol{a_{35}}$
$a_{41} = 1/2$			1	
	$a_{52} = 1/3$			1

Table 7: Comparison a_{35} is required in this matrix.

Step	Independent comparison	Tabu List	Comments
1	a_{32}	-	We search for the first independent comparison with w_3 as a numerator in the row 3.
2	a_{21}	a_{32}; a_{23}	a_{32} and its reciprocal are set in the tabu list, and we search for an independent comparison with w_2 as a numerator in the row 2.
3	a_{14}	a_{32}; a_{23}; a_{21}; a_{12}	a_{21} and its reciprocal are set in the tabu list, and we search for an independent comparison with w_1 as a numerator in the row 1.
4	-	a_{32}; a_{23}; a_{21}; a_{12}; a_{14} ; a_{41}	a_{14} and its reciprocal are set in the tabu list. No independent comparison has w_4 as a numerator and is outside the tabu list; we backtrack to step 2.
5	a_{25}	-	The denominator of a_{25} is equivalent to that of the transitive comparison, and the process stops.

Table 8: Steps in the calculation of the comparison a_{35}.

The final result is: $a_{35} = a_{32} \cdot a_{25} = 2 \cdot 3 = 6$

The research tree corresponding to the five steps in Table 8 is given in Figure 1.

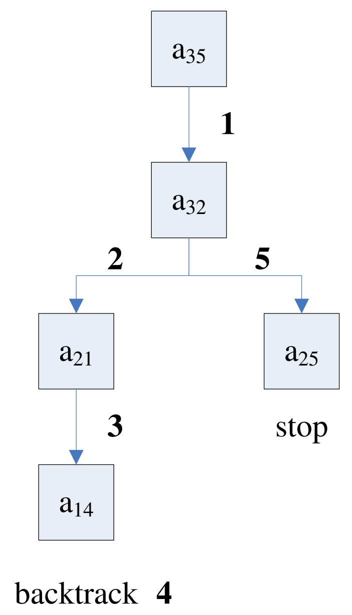


Figure 1: Research tree for a_{35}. The direction of the tree is numbered 1 to 5.

3.4.3 Inconsistency tolerance

For the construction of a fully consistent matrix, only the independent comparisons are required. The other comparisons can be deduced with the transitivity rule (1). The transitive comparisons can be used to assess the accuracy of the independent comparisons. However, our world is not always straight and consistent. A fully consistent matrix cannot be imposed. A degree of inconsistency should be allowed, but also a minimum of consistency should be imposed otherwise a decision could not be taken. Therefore, a tolerance ($\pm\, a_{ij} \cdot t$) is introduced in the relation (5), which gives:

$$a_{ij} \pm a_{ij} \cdot t = a_{ix} \cdot a_{x.} \cdot \ldots \cdot a_{.z} \cdot a_{zj} \qquad (6)$$

where t is a tolerance expressed in percentage.

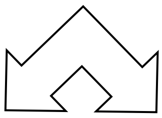
If the extended transitivity rule (6) is violated, the user is required to modify either the value entered (left part of (6)) or the independent comparisons (right part of (6)). Modifying the independent comparisons induces changes to the other previously entered transitive comparisons. The user is offered both matrices and can adopt the most appropriate (Example 4).

Example 4:

We consider the matrix of Table 9. Suppose the user chooses $a_{23} = 5$ and an inconsistency tolerance of $t = 20\%$. According to the transitivity rule, the user is supposed to enter a value $a_{23} = a_{21} \cdot a_{13} = 1/2 \cdot 6 = 3$. This falls outside the tolerance of $5 \pm 5 \cdot 0.2 = [4, 6]$. If the user maintains their entry, the user can modify the independent comparisons and see the effects on other transitive comparisons in a what-if analysis. Then, the user can either adopt the solution proposed by the program (Table 10) or proceed with their own matrix (Table 11).

1	2	6	a_{14}
1/2	1	a_{23}	1/4
1/6	a_{23}	1	a_{34}
a_{41}	4	a_{43}	1

Table 9: Comparison a_{23} is required in this matrix.



1	2	6	a_{14}
1/2	1	**3**	1/4
1/6	**1/3**	1	a_{34}
a_{41}	4	a_{43}	1

Table 10: Matrix with the proposed comparison ($a_{23} = 3$).

1	2	*8*	a_{14}
1/2	1	**5**	1/4
1/2	**1/5**	1	a_{34}
a_{41}	4	a_{43}	1

Table 11: Matrix after changing premises ($a_{13} = 8$).

4. Calculating acceptable inconsistency tolerance

The interactive algorithm proposed in (Ishizaka, Lusti 2004) does not specify the acceptable inconsistency tolerance. In this paper, Monte Carlo simulations have been performed in order to estimate the level of tolerance that should be allowed for a decision maker whilst entering the transitive judgments. This is important in order to limit the CR within the acceptable level (i.e. CR < 0.1). The experiments were performed in Java by repetitive random sampling using the uniform distribution. For each value of n ranging from 3 to 9, a different level of inconsistency tolerance has been applied to reconstruct a positive symmetrically reciprocal matrix (PSRM) from each set of randomly generated independent judgments. PSRMs were constructed by generating a random set of independent judgments and then using Algorithm 1 described in Section 3.4.2 for the dependent judgments. The maximum level of tolerance in (6) was varied from 0% to 80%. For every n, 5,000 PSRMs were generated for each level of tolerance. This makes a total of 31,500 PSRMs generated for this experiment.

The values of tolerance and CR were recorded after each step, and they were used to extract descriptive statistical figures for each value of n. These results are shown in Figure 2 in the form of box plots, where boxes indicate the interquartile range (IQR), vertical lines indicate the range and the horizontal line is the median value. The average (mean) values are connected through a curve with the marker 'x' on each intersection.

For all the randomly generated matrices of order $n = 3$, the values of CR remained below 0.1 when the transitive judgments were allowed the tolerance of 60% or less (Figure 2a). When the tolerance was increased to 70%, some of the matrices became unacceptably inconsistent (CR > 0.1). When raised to 80%, the majority of the matrices became unacceptable in AHP terms. In case of $n = 4$, as shown in Figure 2b, the majority of the matrices became unacceptable at 70% tolerance allowed. So, the tolerance level decreased for $n = 4$. Going further, the level of tolerance keeps decreasing when we increase the value of n. For $n = 9$, the level of tolerance goes below 60%. In other words, as the number of stimuli that need to be compared is increased, the level of tolerance for inconsistent judgments is more likely to be reduced.

It is therefore possible to derive the thresholds of tolerance from the level of inconsistencies found above. Figure 3 highlights the trends of how consistency gets reduced when the level of tolerance is increased. Each curve in the figure depicts Q3 (third quartile) values, implying that 75% of the generated matrices had CR values below this level. Similar results can be calculated with a stricter condition of choosing the highest CR value.

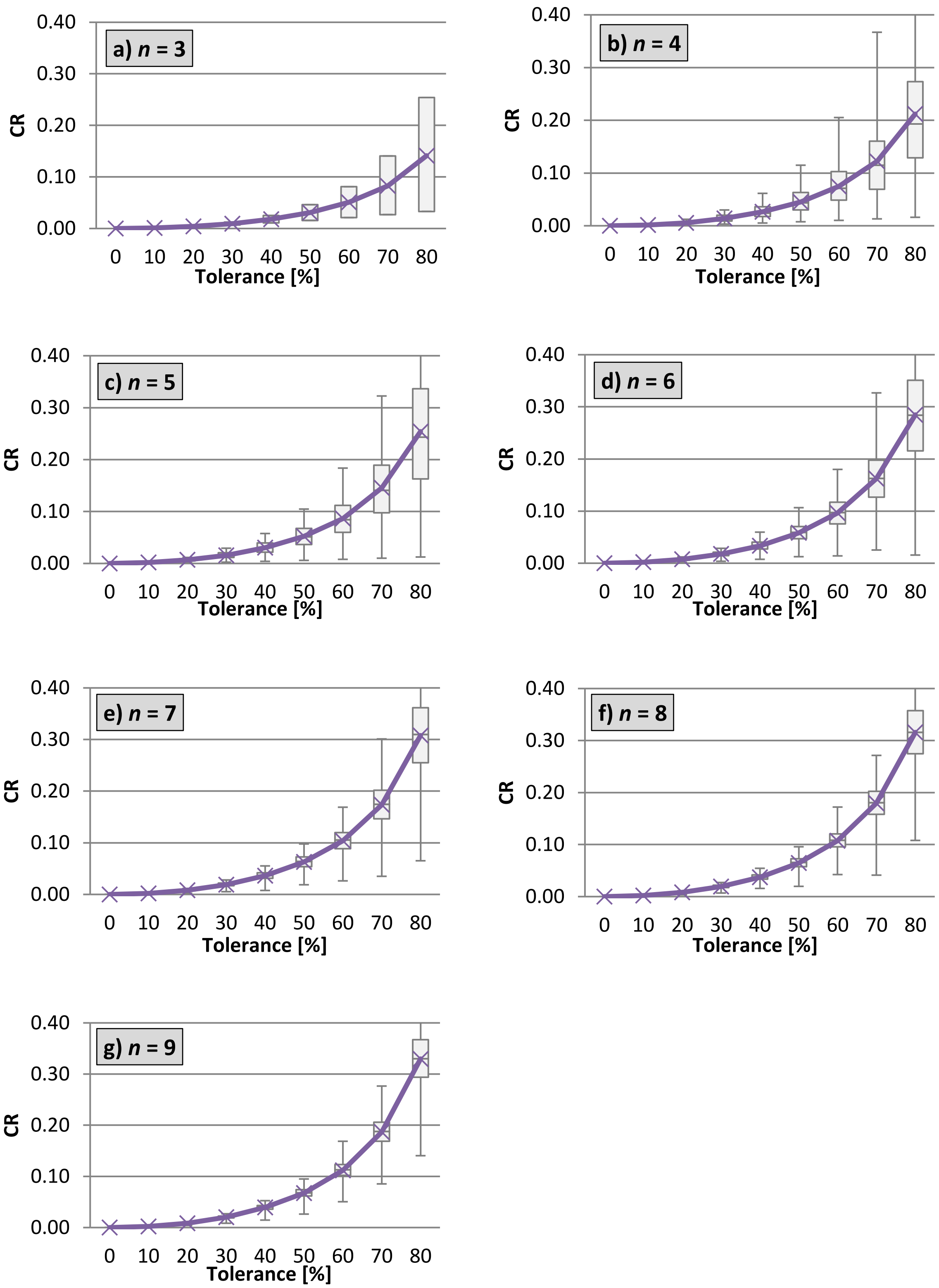


Figure 2: Box plots for the CR values against tolerance. The subplots (a) to (g) represent n from 3 to 9. Lines depict the third quartiles.

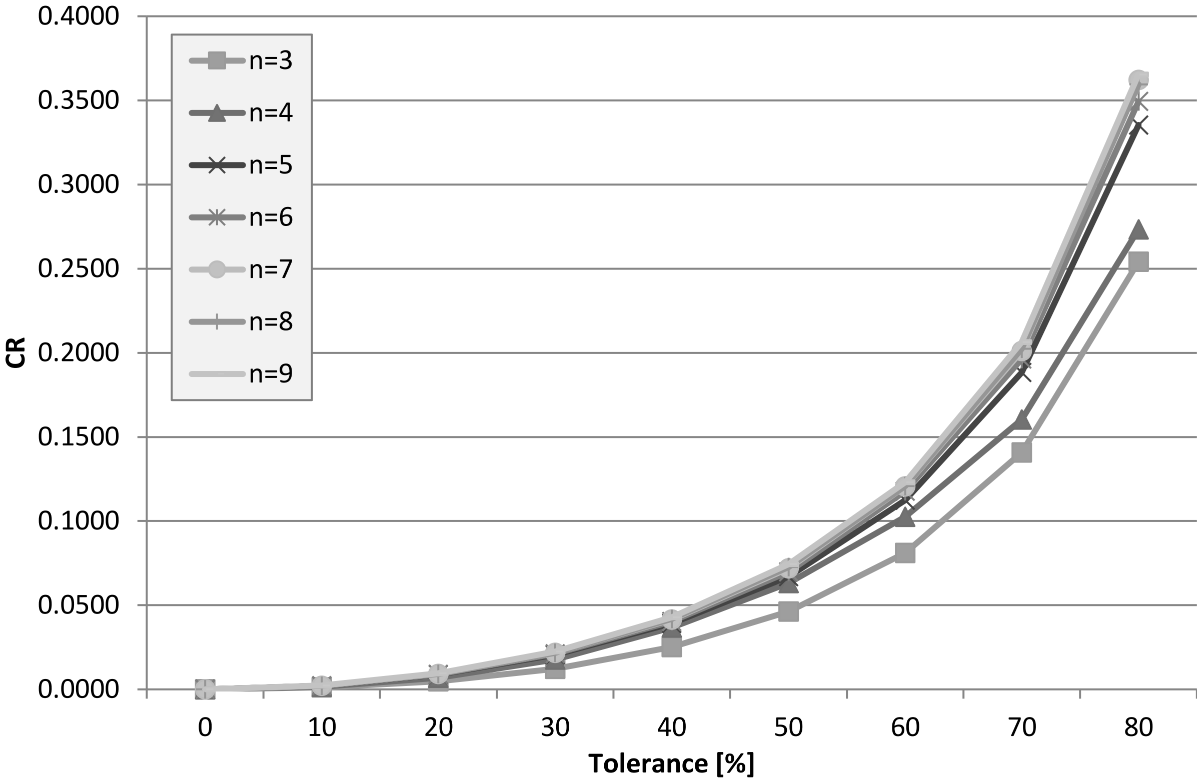


Figure 3: Threshold of tolerance to keep CR below 0.1

For example, looking at the trend for n = 5 (see the curve with the marker 'x' in Figure 3), the value of CR remains below 0.1 when the tolerance is restricted below 57.3%. As soon as the threshold of tolerance is relaxed above 57.3%, the generated comparisons develop a tendency to become unacceptable (CR > 0.1). Therefore, users may be assisted about the allowed tolerance level in an interactive manner when they enter the pairwise comparison judgments.

Similarly, the values of these thresholds can be estimated for each n. The level of thresholds for n = 3 to 9 are derived from these results and are given below (Table 12):

n	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Tolerance	63.2%	59.3%	57.3%	56.3%	55.8%	55.7%	55.3%

Table 12: Tolerance thresholds allowed for n = 3 to 9.

For example, according to Table 12, if the number of stimuli to be compared is 9, the decision maker should not be allowed to deviate more than 55.3% of the judgment value, estimated through the transitivity rule.

It is also worthwhile to note that particular attention is to be given at comparisons of values 1 and 2, because a too high tolerance may induce an inconsistency in direction (i.e. ordinal inconsistency). For $a_{ij} = 1$, the preference equivalence will not be maintained when any level of tolerance is allowed. For $a_{ij} = 2$, a tolerance of 50% is the maximum value that prevents inconsistencies in direction: $2 \pm 2 \cdot 0.5 = [1, 3]$. A higher tolerance would incorporate values lower than 1 and therefore inconsistent in direction. Therefore, it may be appropriate to limit the tolerance threshold to 50% at least for comparisons of value 2. Ordinal inconsistency is not explicitly handled by the consistency check of CR < 0.1 (Siraj et al. 2012), and, therefore, the issue of handling ordinal inconsistencies needs to be investigated separately.

5. Conclusion

AHP is often used to rank alternatives but also for performance measurement. If a pairwise comparison matrix is not consistent enough, ~~then results may be invalid and~~ then the results may suffer from rank reversal and more importantly results may be invalid. ~~In this paper, we have explained that the normalisation of weights and strong local inconsistency are two distinct causes that may produce a rank reversal in AHP. For both causes, we have described the way to avoid a rank reversal. The use of an alternative acting as a reference provides a normalisation of the weights, which is independent of the number of alternatives.~~

The consistency ratio is not a sufficient test because it cannot detect strong local inconsistencies. In this paper, we have described an algorithm that intervenes after each pairwise comparison which contradicts the comparisons made so far, explains it and suggests consistent alternatives. It allows detection of local inconsistencies, which are the cause of rank reversals. This control on the fly is not possible with the other introduced consistency measures (Salo, Hamalainen 1997; Aguarón, Moreno-Jiménez 2003; Golden, Wang 1989; Saaty 1977) because all of them force the decision maker to provide all the comparisons before providing feedback. Consequently, suggestions of changes to resolve inconsistencies could only be discussed at this later point. This is a restriction that does not lend itself to a good interaction between the decision maker and the facilitator of the decision process (Monti, Carenini 1995).

The proposed method has been tested through simulations in order to determine the maximum level of tolerance allowed for the decision maker whilst entering their transitive judgments. In future, the proposed method will be applied in practice, and empirical evidence will be collected in order to gauge the usefulness of this ~~method~~algorithm.

Appendix: Analytic hierarchy process

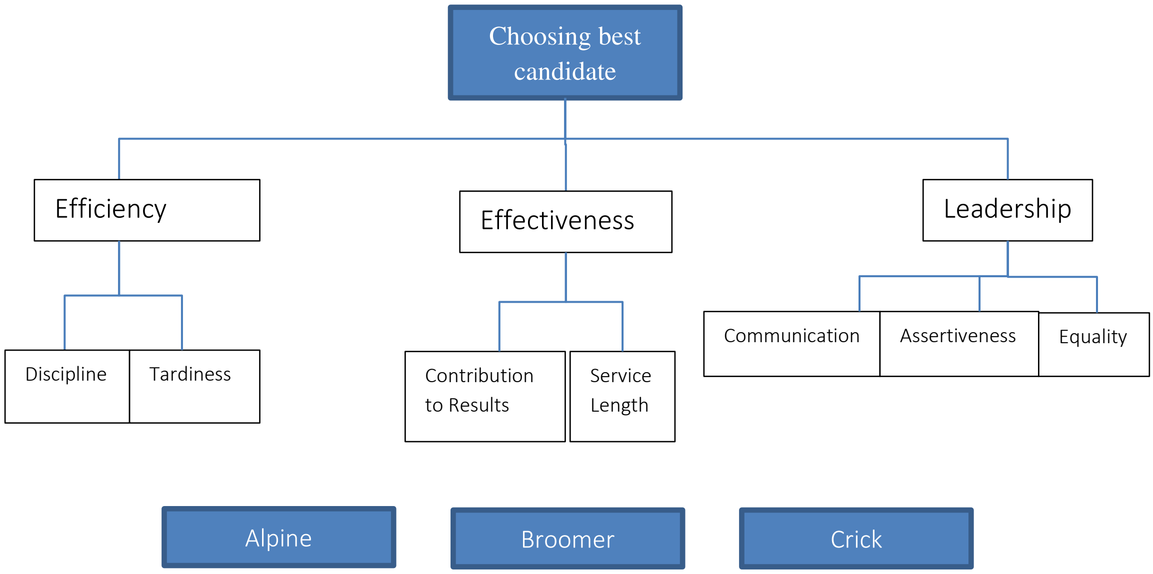
In the analytic hierarchy process, a decision maker can model a given decision problem using a hierarchical structure. The overall goal is at the root of this hierarchy, and each node represents a criterion or sub-criterion that is important to consider in attaining the overall goal. Alternatives are placed at the lowest level of this hierarchy.

For example, let us consider an imaginary performance analysis exercise of evaluating staff for an award in an organisation. The panel can use the analytic hierarchy process to make this analysis in a structured way. They sit together and create a hierarchy of criteria that are important for this performance analysis, and they come up with something as follows:

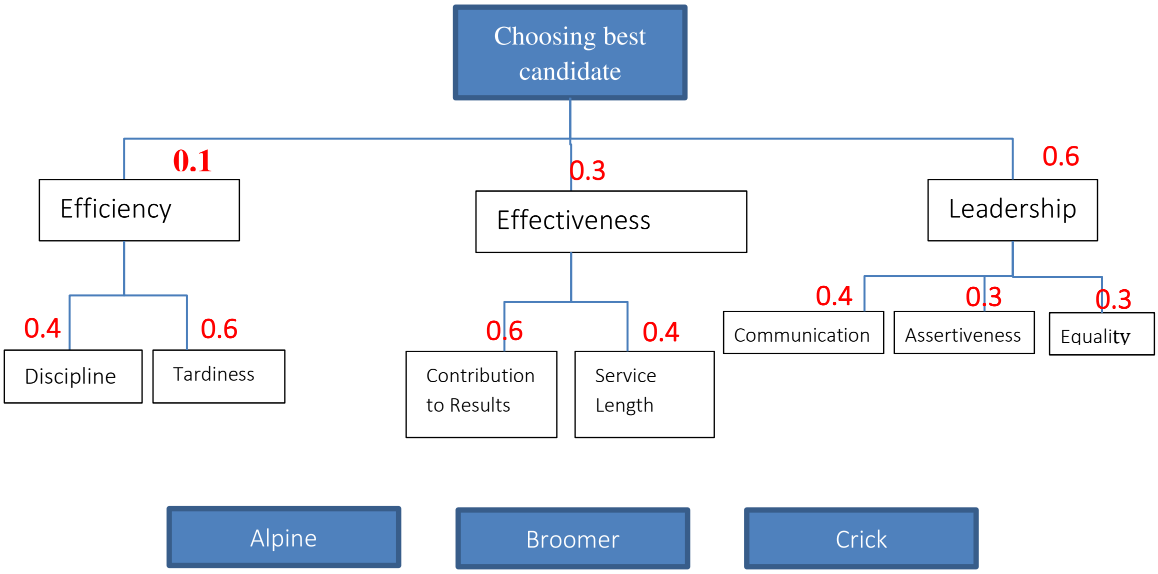
- Choosing a candidate for best performance award
    - Efficiency
        - Discipline
        - Tardiness
    - Effectiveness
        - Contribution to Results
        - Service Length
    - Leadership
        - Communication

- Assertiveness
- Equality

After receiving the nominations, the reviewing panel has shortlisted three candidates to pick one from, and their names are: Alpine, Broomer, and Crick. The hierarchical representation of this problem can be demonstrated in the following diagram:



The preferences of the decision maker are obtained through pairwise comparisons, that is, we ask the decision maker to compare each pair of options (or criteria) and tell which option (or criterion) is more preferred and how much. For example, a decision maker may say that the criterion of Effectiveness is twice as important as the criterion of Efficiency, and the criterion of Leadership is three times more important than Effectiveness. The decision maker might also add that the criterion of Leadership is six times more important than the Efficiency. These pairwise comparisons are then used to calculate the preference weights for each criterion and option. For example, in this case, we can elicit the preference weights of 60%, 30% and 10% for Leadership, Effectiveness and Efficiency, respectively. These percentages can also be represented in terms of fractions that add up to 1, that are, 0.6, 0.3 and 0.1. The same exercise can be done for all criteria and alternatives in order to obtain these preference weights at each level of the hierarchy. An example is provided below in the diagram.



So far, we assumed that the pairwise comparison judgments are internally consistent with each other, that is, they follow the mathematical rule of transitivity. For example, if Effectiveness is 3 times more important than Efficiency, and Leadership is twice more important than Effectiveness, then mathematically, Leadership must be 6 times more important than Efficiency. However, in practice, this rule is not strictly followed by people (this is widely discussed and debated in the decision-making literature). Eliciting weights from these inconsistent matrices is not a straightforward calculation. In such situations, the traditional method to calculate the preference weights is to use the principal eigenvector method, originally used by Saaty for the pairwise comparison matrices. The computation of the eigenvector involves algebraic equations. Another widely used approach is to calculate the geometric mean of all the elements (a_{ij}) in each row (represented by the notation i) of the matrix. The mathematical formulation is simpler than the eigenvector method, i.e.

$$r_i = \left(\prod_j a_{ij}\right)^{\frac{1}{n}}$$

From a usefulness perspective, the calculation of an eigenvector demands more effort in Microsoft Excel while the operation of the geometric mean is as simple as taking average of numbers.

It is very important to assess the level of inconsistency in judgments, as inconsistent judgments might be misleading for eliciting preferences. If the level of inconsistency is too high, then the whole process of eliciting preferences may be questioned. Saaty, the inventor of AHP, proposed to use the maximal eigenvalue (a value that relates to the principal eigenvector) in order to assess the level of inconsistency. Based on this, he proposed a measure called the consistency ratio (or CR) along with a threshold value of 0.1 where any set of judgments becomes ~~unacceptable~~unreliable (or questionable) if the CR value exceeds this threshold. When the CR is less than 0.1, it is considered safe to elicit the preference weights. For example, the decision makers may obtain the following table of information containing individual scores obtained by each candidate in each criterion:

	Efficiency 10%		**Effectiveness 30%**		**Leadership 60%**		
	Discipline 40%	Tardiness 60%	Contribution to Results 60%	Service Length 40%	Communication 40%	Assertiveness 30%	Equality 30%
Alpine	0.4	0.4	0.3	0.2	0.3	0.1	0.2
Broomer	0.3	0.2	0.3	0.4	0.4	0.6	0.4
Crick	0.3	0.4	0.4	0.4	0.3	0.3	0.4

Note that the column headers show the weight of importance for each criterion at each level of hierarchy as well. This whole table of information is sometimes referred to as the decision table.

Once all the scores and preference weights are calculated (or estimated), the final score for each alternative is aggregated with the help of weighted sum approach.

	Efficiency 0.1		Effectiveness 0.3		Leadership 0.6			
	Discipline 0.4*0.1	Tardiness 0.6*0.1	Contribution to Results 0.6*0.3	Service Length 0.4*0.3	Communication 0.4*0.6	Assertiveness 0.3*0.6	Equality 0.3*0.6	**Weighted Sum**
Alpine	0.4	0.4	0.3	0.2	0.3	0.1	0.2	**0.244**

Broomer	0.3	0.2	0.3	0.4	0.4	0.6	0.4	**0.402**
Crick	0.3	0.4	0.4	0.4	0.3	0.3	0.4	**0.354**

As we can see, the scores calculated using the weighted sum method suggest that the second candidate called Broomer has the highest overall score, followed by Crick who obtained 0.354, and the least favourite turns out to be Alpine with the score of 0.244.

Although the AHP method assigns a quantitative score (or value) in the end, one must carefully interpret these scores as they are based on subjective judgments and elicitation methods that may involve inconsistent judgments. Therefore, it is usually good practice to check and discuss the robustness of these final scores by changing weights of criteria (and/or individual scores of alternatives). If the final scores do not change drastically when we make small changes in these weights, we can say that these scores are robust, and taking the decision based on this analysis might be justified.

References

Aguarón, J., Moreno-Jiménez, J.: The Geometric Consistency Index: Approximated Thresholds. European Journal of Operational Research **147**(1), 137-145 (2003)

Bana e Costa, C., Vansnick, J.: A Critical Analysis of the Eigenvalue Method Used to Derive Priorities in AHP. European Journal of Operational Research **187**(3), 1422-1428 (2008)

Barzilai, J., Lootsma, F.: Power relation and group aggregation in the multiplicative AHP and SMART. Journal of Multi-Criteria Decision Analysis **6**(3), 155-165 (1997)

Belton, V., Gear, A.: On a Shortcoming of Saaty's Method of Analytical Hierarchies. Omega **11**(3), 228-230 (1983)

Brunelli, M.: A survey of inconsistency indices for pairwise comparisons. International Journal of General Systems **47**(8), 751-771 (2018). doi:10.1080/03081079.2018.1523156

Brunelli, M., Canal, L., Fedrizzi, M.: Inconsistency indices for pairwise comparison matrices: a numerical study. Annals of Operations Research **211**(1), 493-509 (2013). doi:10.1007/s10479-013-1329-0

Camerer, C.: Individual decision making. In: Kagel, J., Roth, A. (eds.) The Handbook of Experimental Economics. pp. 587-703. Princeton University Press, Princeton (1995)

Cao, D., Leung, L.C., Law, J.S.: Modifying Inconsistent Comparison Matrix in Analytic Hierarchy Process: a Heuristic Approach. Decision Support Systems **44**(4), 944-953 (2008)

Corbin, R., Marley, A.A.J.: Random utility models with equality: an apparent, but not actual, generalization of random utility models. Journal of Mathematical Psychology **11**(3), 274-293 (1974)

De Keyser, W., Peeters, P.: A Note on the Use of PROMETHEE Multicriteria Methods. European Journal of Operational Research **89**(3), 457-461 (1996)

Dyer, J.: A clarification of "Remarks on the Analytic Hierarchy Process". Management Science **36**(3), 274-275 (1990a)

Dyer, J.: Remarks on the Analytic Hierarchy Process. Management Science **36**(3), 249-258 (1990b)

Ergu, D., Kou, G., Peng, Y., Shi, Y.: A simple method to improve the consistency ratio of the pair-wise comparison matrix in ANP. European Journal of Operational Research **213**(1), 246-259 (2011). doi:http://dx.doi.org/10.1016/j.ejor.2011.03.014

Ergu, D., Kou, G., Shi, Y., Shi, Y.: Analytic network process in risk assessment and decision analysis. Computers & Operations Research **advance online publication** **http://dx.doi.org/10.1016/j.cor.2011.03.005** (2013)

Figueira, J.R., Roy, B.: A note on the paper, "Ranking irregularities when evaluating alternatives by using some ELECTRE methods", by Wang and Triantaphyllou, Omega (2008). Omega **37**(3), 731-733 (2009)

Finan, J.S., Hurley, W.J.: The Analytic Hierarchy Process: Can Wash Criteria Be Ignored? . Computers and Operations Research **29**(8), 1025-1030 (2002)

Forman, E., Gass, S.: The Analytic Hierarchy Process – An Exposition. Operations Research **49**(4), 469-486 (2001)

García-Cascales, S., Lamata, T.: On rank reversal and TOPSIS method. Mathematical and Computer Modelling **56**(5–6), 123-132 (2012). doi:http://dx.doi.org/10.1016/j.mcm.2011.12.022

Gaul, W., Gastes, D.: A note on consistency improvements of AHP paired comparison data. Advances in Data Analysis and Classification **6**(4), 289-302 (2012). doi:10.1007/s11634-012-0119-x

Golden, B., Wang, Q.: An Alternate Measure of Consistency. In: Golden, B., Wasil, E., Harker, P. (eds.) The Analytic Hierarchy Process: Application and Studies pp. 68–81. Springer-Verlag, New-York (1989)

González-Pachón, J., Romero, C.: A method for dealing with inconsistencies in pairwise comparisons. European Journal of Operational Research **158**(2), 351-361 (2004)

Harker, P., Vargas, L.: The Theory of Ratio Scale Estimation: Saaty's Analytic Hierarchy Process. Management Science **33**(11), 1383-1403 (1987)

Harker, P., Vargas, L.: Reply to “Remarks on the Analytic Hierarchy Process”. Management Science **36**(3), 269-273 (1990)

Holder, R.: Some Comment on the Analytic Hierarchy Process. Journal of the Operational Research Society **41**(11), 1073-1076 (1990)

Holder, R.: Response to Holder's Comments on the Analytic Hierarchy Process: Response to the Response. Journal of the Operational Research Society **42**(10), 914-918 (1991)

Ishizaka, A., Balkenborg, D., Kaplan, T.: Influence of aggregation and measurement scale on ranking a compromise alternative in AHP. Journal of the Operational Research Society **62**(4), 700-710 (2010)

Ishizaka, A., Labib, A.: Review of the main developments in the analytic hierarchy process. Expert Systems with Applications **38**(11), 14336-14345 (2011a). doi:DOI 10.1016/j.eswa.2011.04.143

Ishizaka, A., Labib, A.: Selection of new production facilities with the Group Analytic Hierarchy Process Ordering method. Expert Systems with Applications **38**(6), 7317–7325 (2011b). doi:DOI: 10.1016/j.eswa.2010.12.004

Ishizaka, A., López, C.: Cost-benefit AHPSort for performance analysis of offshore providers. International Journal of Production Research, online advance production, doi: 10.1080/00207543.00202018.01509393 (2019). doi:10.1080/00207543.2018.1509393

Ishizaka, A., Lusti, M.: An Expert Module to Improve the Consistency of AHP Matrices. International Transactions in Operational Research **11**(1), 97-105 (2004)

Kendall, M.G., Smith, B.B.: On the Method of Paired Comparisons. Biometrika **31**(3/4), 324-345 (1940)

Keskin, B., Köksal, C.: A hybrid AHP/DEA-AR model for measuring and comparing the efficiency of airports. International Journal of Productivity and Performance Management **68**(3), 524-541 (2019). doi:10.1108/IJPPM-02-2018-0043

Kwiesielewicz, M., van Uden, E.: Inconsistent and Contradictory Judgements in Pairwise Comparison Method in AHP. Computers and Operations Research **31**(5), 713-719 (2004)

Linares, P.: Are inconsistent decisions better? An experiment with pairwise comparisons. European Journal of Operational Research **193**(2), 492-498 (2009)

Lootsma, F.: Scale sensitivity in the multiplicative AHP and SMART. Journal of Multi-Criteria Decision Analysis **2**(2), 87-110 (1993)

Maleki, H., Zahir, S.: A Comprehensive Literature Review of the Rank Reversal Phenomenon in the Analytic Hierarchy Process. Journal of Multi-Criteria Decision Analysis, n/a-n/a (2012). doi:10.1002/mcda.1479

Mareschal, B., De Smet, Y., Nemery, P.: Rank Reversal in the PROMETHEE II Method : Some New Results. Paper presented at the IEEE 2008 International Conference on Industrial Engineering and Engineering Management, Singapore,

Millet, I., Saaty, T.: On the Relativity of Relative Measures-Accommodating both Rank Preservation and Rank Reversals in the AHP. European Journal of Operational Research **121**(1), 205-212 (2000)

Monti, S., Carenini, G.: Dealing with the Expert Inconsistencies: the Sooner the Better. Paper presented at the IJCAI-95 Workshop: "Building Probabilistic Networks: where do the numbers come from?", Montreal, Perez, J, Jimeno, J.L., Mokotoff, E.: Another Potential Shortcoming of AHP. TOP **14**(1), 99-111 (2006)

Pérez, J.: Some comments on Saaty's AHP. Management Science **41**(6), 1091-1095 (1995)

Saaty, T.: A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology **15**(3), 234-281 (1977)

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

Saaty, T.: Rank Generation, Preservation and Reversal in the Analytic Hierarchy Decision Process. Decision Sciences **18**(2), 157-177 (1987)

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

Saaty, T.: Response to Holder's Comments on the Analytic Hierarchy Process. Journal of the Operational Research Society **42**(10), 909-929 (1991)

Saaty, T.: Decision-making with the AHP: Why is the Principal Eigenvector necessary? European Journal of Operational Research **145**(1), 85-91 (2003)

Saaty, T.: Rank from Comparisons and from Ratings in the Analytic Hierarchy/Network Processes. European Journal of Operational Research **168**(2), 557-570 (2006)

Saaty, T., Sagir, M.: An Essay on Rank Preservation and Reversal. Mathematical and Computer Modelling **advance online publication**, doi:10.1016/j.mcm.2008.1008.1001 (2008)

Saaty, T., Takizawa, M.: Dependence and Independence: from Linear Hierarchies to Nonlinear Networks. European Journal of Operational Research **26**(2), 229-237 (1986)

Saaty, T., Vargas, L.: The legitimacy of rank reversal. Omega **12**(5), 513-516 (1984)

Saaty, T., Vargas, L.: The Analytic Hierarchy Process: wash criteria should not be ignored. International Journal of Management and Decision Making **7**(2/3), 180-188 (2006)

Salo, A., Hamalainen, R.: On the Measurement of Preference in the Analytic Hierarchy Process. Journal of Multi-Criteria Decision Analysis **6**(6), 309-319 (1997)

Schoner, B., Wedley, W.: Ambiguous Criteria Weights in AHP: Consequences and Solutions. Decision Sciences **20**(3), 462-475 (1989)

Schoner, B., Wedley, W., Choo, E.: A Unified Approach to AHP with Linking Pins. European Journal of Operational Research **64**(3), 384-392 (1993)

Schoner, B., Wedley, W.C., Choo, E.U.: A Rejoinder to Forman on AHP, with Emphasis on the Requirements of Composite Ratio Scales. Decision Sciences **23**(2), 509–517 (1992)

Sipahi, S., Timor, M.: The analytic hierarchy process and analytic network process: an overview of applications. Management Decision **48**(5), 775-808 (2010)

Siraj, S., Mikhailov, L., Keane, J.: A heuristic method to rectify intransitive judgments in pairwise comparison matrices. European Journal of Operational Research **216**(2), 420-428 (2012). doi:http://dx.doi.org/10.1016/j.ejor.2011.07.034

Subramanian, N., Ramanathan, R.: A review of applications of Analytic Hierarchy Process in operations management. International Journal of Production Economics **138**(2), 215-241 (2012). doi:http://dx.doi.org/10.1016/j.ijpe.2012.03.036

Tone, K.: A comparative study of AHP and DEA. Theory and applications. In: Tone, K. (ed.) Advances in DEA Theory and Applications. Wiley, Chichester (2017)

Triantaphyllou, E.: Two new cases of rank reversals when the AHP and some of its additive variants are used that do not occur with the Multiplicative AHP. Journal of Multi-Criteria Decision Analysis **10**(1), 11-25 (2001)

Troutt, M.: Rank Reversal and the Dependence of Priorities on the Underlying MAV Function. Omega **16**(4), 365-367 (1988)

Tversky, A., Slovic, P., Kahneman, D.: The cause of preference reversal. American Economic Association **80**(1), 204-217 (1990)

Vargas, L.: Comments on Barzilai and Lootsma Why the Multiplicative AHP is Invalid: A Practical Counterexample. Journal of Multi-Criteria Decision Analysis **6**(4), 169-170 (1997)

Veni, K., Rajesh, R., Pugazhendhi, S.: Development of decision making model using integrated AHP and DEA for vendor selection. Procedia Engineering **38**, 3700-3708 (2012). doi:https://doi.org/10.1016/j.proeng.2012.06.425

Wang, Y., Chin, K.-S., Luo, Y.: Aggregation of direct and indirect judgements in pairwise comparison matrices with a re-examination of the criticisms by Bana e Costa and Vansnick. Information Sciences **179**(3), 329-337 (2009)

Wang, Y., Elhag, T.: An Approach to Avoiding Rank Reversal in AHP. Decision Support Systems **42**(3), 1474-1480 (2006)

Wang, Y., Luo, Y.: On Rank Reversal in Decision Analysis. Mathematical and Computer Modelling **49**(5-6), 1221-1229 (2009)

Wanga, X., Triantaphyllou, E.: Ranking Irregularities when Evaluating Alternatives by Using some ELECTRE Methods. Omega **36**(1), 45-63 (2008)

Xu, Z., Wei, C.: A Consistency Improving Method in the Analytic Hierarchy Process. European Journal of Operational Research **116**(2), 443-449 (1999)