Are inconsistent decisions better? A follow-up interactive experiment with pairwise comparisons

ABSTRACT

In a previous experimental study, Linares (2009) observed that an automatic correction of inconsistency worsens the preference representation of the decision maker. In this paper, the experimental study is repeated but with an automatic and an interactive approach to correct the inconsistencies. Two new problems are used: one with objective (tangible) alternatives, the other with subjective (intangible) alternatives. The new results confirm the previous observation and further find that the interactive method does not better represent decision makers' preferences. Furtheremore, it was also observed in the objective problem, that the original priorities were closer to the true values than those interactively or automatically corrected, making the interactive effort to reduce inconsistencies questionable.

Keywords: AHP, pairwise comparisons, consistency, automatic correction, interactive correction.

1. Introduction

Inconsitencies in pairwise comparisons arise when the transitity or reciprocity rules are not respected. A violated reciprocity rule is easy to detect as it only involves two judgements at reciprocal positions in a matrix. A broken transitivity rule is more difficult to identify because it involves at least three judgements at any position in the matrix. Indices have been developed to measure the rate of this inconsistency (Brunelli, Canal, & Fedrizzi, 2013; Brunelli & Fedrizzi, 2014; Lin, Kou, & Ergu, 2013). It has been long suggested that a matrix should not exceed 10% inconsistency to have meaningful results (Saaty, 1977; T. Saaty, 1980). A higher inconsistency indicates an illogical assesment or an error in the input due to either the complexity of the decision problem or the limited capability or skills of the decision-maker. There are two ways of correcting inconsistencies:

- An *interactive process* with the decision-maker. The most inconsistent entry is highlighted and the decision-maker asked to reassess it. This process is repeated with the next inconsistent entry, and so on, until the matrix consistency is below the acceptable threshold (T. Saaty, 2003).

- An *automatic process*. The decision-maker is not involved in the correction. It is a technique that is often used when the participants are not available following the data collection, for example, after a survey, to modify the entries. Several automatic corrective methods have been proposed (Benítez, Izquierdo, Pérez-García, & Ramos-Martínez, 2014; Cao, Leung, & Law, 2008; Ergu, Kou, Peng, & Shi, 2011; Jacinto González-Pachón & Romero, 2004; Kou, Ergu, & Shang, 2014; Ma & Li, 2011; Siraj, Mikhailov, & Keane, 2012; Xu & Xia, 2014; Zhang, Sekhari, Ouzrout, & Bouras, 2014).

Karapetrovic and Rosenbloom (1999) were the first to challenge the necessity of improving consistency. They presented three examples of matrices, where in their opinion a correction would not be necessary. Then, Linares (2009) used experimental methods for testing the automatic correction. He asked 18 male graduate engineer students to pairwise compare five different compact cars globally and in term of aesthetics. Inconsistent matrices were automatically corrected and the participants were asked, both before and after correction, which output they preferred. In both cases, the majority preferred the initial outcome. One explanation of this deterioration of the preference elicitation through the automatic correction, is given by Gaul and Gastes (2012) in a Monte-Carlo simulation. They first constructed perfectly consistent matrices, introduced some noise and finally used three automatic corrective methods to improve the consistency. With all three methods, they found that the initial priorities were closer to the priorities from the inconsistent matrices than to those calculated from the automatically corrected matrices. If it is clear that an automatic consistency correction is deteriorating the outcome, there is no shown evidence on the benefits of an interactive correction. As further research (Linares, 2009) states: "*It would be interesting to replicate this experiment when the improvements in consistency are achieved through an interaction with the decision maker, as Saaty (1980) proposed originally*" (p. 497). This is the aim of this paper. Sixty-two participants were asked to solve two problems: one with objective measures, the other with subjective measures. These entries were corrected with an interactive approach and in parallel automatic correction was used. The priorities from the inconsistent matrices were significantly preferred over those of both corrected matrices and in the objective problem, the original priorities were closer to the true values than either of the interactively or automatically corrected.

The structure of this paper is as follows: Section 2 describes the experiments. Section 3 presents and discusses the results and Section 4 concludes the paper.

2. Description of the experiments

2.1 Problem description

Each participant was asked to solve two problems, one with objective and the other with subjective criterion:

a) *Decision with subjective criterion*: Participants were asked to pairwise compare their preferred mode of transportation for a journey from Portsmouth to Gatwick Airport. They were given five alternatives: Train, Coach, Taxi, Car sharing and Own car.

b) *Decision with objective criterion*: Participants were asked to evaluate which of the five cities is closest to Portsmouth: Cardiff, London, Edinburgh, Southampton, and Liverpool.

Both experiments followed the procedure described in the next section.

2.2 Experimental procedure

On average, the experiment lasted half an hour and produced three rankings for each participant:

a) *Original Ranking*: the priorities calculated before inconsistency correction.

b) *Automatic Ranking*: the priorities calculated after inconsistencies were automatically corrected with the goal programming method (Jacinto González-Pachón & Romero, 2004).

c) *Interactive Ranking*: the priorities calculated after the inconsistencies were interactively corrected (T. L. Saaty, 1980).

In the three cases, the priorities are calculated with the eigenvalue method (T. L. Saaty, 1980). Others methods of priorities derivation exists (J González-Pachón & Romero, 2007; Ishizaka & Labib, 2011) but are less used in practice. The procedure is divided into 6 steps:

1. The decision problem (section 2.1) is explained to the participant.

2. The participant pairwise compares the five alternatives. The Consistency Ratio (1), the Original Ranking and the Automatic Ranking are calculated.

$$CR = CI/RI, \tag{1}$$

where: RI is the random index (the average CI of 500 randomly filled matrices)

$$CI = \frac{\lambda_{max} - n}{n-1}, \tag{2}$$

where: n = dimension of the matrix

λ_{max} = maximal eigenvalue

3. If the consistency ratio is acceptable, i.e. below 10%, the experiment terminates otherwise the consistency error ε_{ij} of each pairwise comparison, is calculated with (T. L. Saaty, 2003):

$$\varepsilon_{ij} = \max\left(a_{ij} * \frac{p_j}{p_i}, a_{ji} * \frac{p_i}{p_j}\right) \tag{3}$$

   where: a_{ij} is the pairwise comparison between alternative i with j and p_i, p_j are their respective priorities.

4. The participant is given the possibility to revise the most inconsistent comparison, i.e. the comparison with the highest ε_{ij}. If they decline revising it, they are asked if they want to revise the next most inconsistent comparison. When they revise a comparison, the process restarts from point 3 until the inconsistency falls below 10% or the participant has considered all entries. It is noted that the participant is not forced to revise the comparisons and could keep the same initial consistent matrix if they wish.

5. The final Interactive Ranking is calculated.

6. The participant is asked which of the three rankings (without knowing how they have been calculated) represents their preference.

All data is processed in Excel, where a macro has been programmed to calculate the priorities with the eigenvalue method. An embedded link to the software LINGO has also been added to automatically correct the inconsistencies with the goal programming method.

3. Results

Sixty-two participants took part in the experiment. The first thirty-one participants solved the problem with the subjective criterion first and then the problem with the objective criterion. The next thirty-one participants solved the problems in the reverse order. As already observed in (Bozóki, Dezső, Poesz, & Temesi, 2013), the order did not have any influence on the results. Both samples produced statistically identical outcomes. The collected data can be found in the supplementary files.

3.1 Consistency improvement with the interactive method

When using the interactive consistency correction method, it is possible to calculate the new comparison to make the matrix consistent (T. Saaty, 2003). This information was not communicated as it may have influenced the participants. The only suggestion was to revise the comparisons from most to least inconsistent. The participant could refuse to revise the comparisons. Table 1 shows the number of matrices where the consistency was improved among those originally over 10% inconsistency (participants were not asked to revise matrices below 10% inconsistency). In 100% of the cases, the interactive method improved the consistency.

	Final matrix with improved consistency	Final matrix consistency improved not improved
Problem with subjective criterion	39	0
Problem with objective criterion	34	0

Table 1: Number of matrices where inconsistency improved

3.2 Alternatives ranking

All three rankings ordered the distance of the cities to Portsmouth correctly (Table 2), however, the Automatic ranking is furthest from the true distance (Table 3). The original and interactive priorities are very close, which makes the effort to improve the consistency questionable, if the final result does not improve.

City	Original Ranking	Interactive Ranking	Automatic Ranking	Normalised true distance
Cardiff	0.186 ± 0.069	0.190 ± 0.069	0.226 ± 0.081	0.162
London	0.260 ± 0.053	0.261 ± 0.053	0.267 ± 0.058	0.267
Edinburgh	0.046 ± 0.040	0.046 ± 0.041	0.068 ± 0.056	0.022
Southampton	0.423 ± 0.104	0.419 ± 0.093	0.338 ± 0.087	0.471
Liverpool	0.085 ± 0.033	0.084 ± 0.032	0.101 ± 0.036	0.078

Table 2: Closest city to Portsmouth

City	|Original priority - true distance|	|Interactive priority - true distance|	|Automatic priority - true distance|
Cardiff	0.087	0.082	0.140
London	0.057	0.058	0.089
Edinburgh	0.042	0.040	0.048
Southampton	0.024	0.024	0.046
Liverpool	0.023	0.022	0.029

Table 3: Error between estimated and true distances

As expected, the priorities of the subjective problems were more dispersed, i.e. the standard deviation was higher (Table 4) than in the objective problem (Table 2). The “own car” alternative was by far the most preferred transportation mode in the original and interactive ranking. This clear preference for ‘own car’ was faded in the interactive ranking. As there were no true values in this problem, the participants were asked which ranking they preferred to determine if a consistency improvement was useful. The next section analyses this.

City	Original priority	Interactive priority	Automatic priority
Train	0.187 ± 0.131	0.190 ± 0.132	0.183 ± 0.123
Coach	0.121 ± 0.104	0.127 ± 0.099	0.131 ± 0.098
Taxi	0.179 ± 0.142	0.173 ± 0.145	0.167 ± 0.133
Car sharing	0.176 ± 0.134	0.181 ± 0.137	0.161 ± 0.105
Own car	0.337 ± 0.219	0.214 ± 0.230	0.357 ± 0.229

Table 4: Priorities of transport selection

3.3 Participants' preferred ranking

In the case of the subjective problem, all rankings were equally preferred (Table 5). A Chi-square test confirms that the frequencies of participants' preferences were equally distributed.

	Participants' preference of rankings		
	Interactive	Automatic	Original
Observed Frequency	13	14	12
Expected Frequency (proportion)	13 $(.\overline{3})$	13 $(.\overline{3})$	13 $(.\overline{3})$

Note. $\chi 2 = 0.15$, degree of freedom =2, significance threshold $p > .05$

Table 5: Frequencies of participants' preferred ranking for the subjective problem

For the objective problem, the automatic consistency correction was by far the least preferred ranking (Table 6). A Chi-square test confirmed that the frequency of the participants' preferences were not equally distributed. If the automatic ranking is ignored, there is no significant difference between the original and interactive ranking with a Chi-square test.

	Participants' preference of rankings		
	Interactive	Automatic	Original
Observed Frequency	18	4	12
Expected Frequency (proportion)	$11.\overline{3}$ $(.\overline{3})$	$11.\overline{3}$ $(.\overline{3})$	$11.\overline{3}$ $(.\overline{3})$

Note. $\chi 2 = 8.71$, degree of freedom =2, significance threshold $p < .05$

Table 6: Frequency of participants' preferred ranking for the objective problem

3.4 Closest ranking to the true value

For the objective problem, the ranking closest to the true value is calculated (Table 7). The original ranking is more often closest to the true values. The Chi-square test confirms that frequencies are not equally distributed.

	Rankings closest to the true value		
	Interactive	Automatic	Original
Observed Frequency	6	7	18
Expected Frequency (proportion)	$10.\overline{3}\ (.\overline{3})$	$10.\overline{3}\ (.\overline{3})$	$10.\overline{3}\ (.\overline{3})$

Note. $\chi 2$ = 8.58*, degree of freedom =2, significance threshold $p < .05$

Table 7: Rankings closest to the true value

4. Conclusions

In this paper, two decision problem experiments were conducted where inconsistent matrices were corrected automatically and interactively. Here it was shown that the interactive and automatic methods improved consistencies in pairwise comparisons. It also confirmed the results of the study by (Linares, 2009), which found that the automatic correction of inconsistencies does not better represent the preferences of decision-makers. More importantly it was shown that an interactive correction does not better represent the preferences of decision-makers and lastly, the original ranking is closest to the true value in the objective problem.

The difference between the priorities of the original and interactive ranking were found to be very small. No rank reversal was observed following the interactive corrections, therefore, it is possible that the participants experienced some difficulty when choosing their preferred ranking. Removing intransitivities may be aesthetically appealing, but the benefits are unclear. This corrective effort may only be justified for high inconsistent matrices and consequently, the 10% consistency threshold should be immensely increased. This hypothesis provides an investigation for future research.

REFERENCES

Benítez, J., Izquierdo, J., Pérez-García, R., & Ramos-Martínez, E. (2014). A simple formula to find the closest consistent matrix to a reciprocal matrix. *Applied Mathematical Modelling, 38,* 3968-3974.

Bozóki, S., Dezső, L., Poesz, A., & Temesi, J. (2013). Analysis of pairwise comparison matrices: an empirical research. *Annals of Operations Research, 211*, 511-528.

Brunelli, M., Canal, L., & Fedrizzi, M. (2013). Inconsistency indices for pairwise comparison matrices: a numerical study. *Annals of Operations Research, 211*, 493-509.

Brunelli, M., & Fedrizzi, M. (2014). Axiomatic properties of inconsistency indices for pairwise comparisons. *Journal of the Operational Research Society, 66*, 1-15.
Cao, D., Leung, L. C., & Law, J. S. (2008). Modifying Inconsistent Comparison Matrix in Analytic Hierarchy Process: a Heuristic Approach. *Decision Support Systems 44*, 944-953.
Ergu, D., Kou, G., Peng, Y., & Shi, Y. (2011). A simple method to improve the consistency ratio of the pair-wise comparison matrix in ANP. *European Journal of Operational Research, 213*, 246-259.
Gaul, W., & Gastes, D. (2012). A note on consistency improvements of AHP paired comparison data. *Advances in Data Analysis and Classification, 6,* 289-302.
González-Pachón, J., & Romero, C. (2004). A method for dealing with inconsistencies in pairwise comparisons. *European Journal of Operational Research, 158*, 351-361.
González-Pachón, J., & Romero, C. (2007). Inferring consensus weights from pairwise comparison matrices without suitable properties. *Annals of Operations Research, 154*, 123-132.
Ishizaka, A., & Labib, A. (2011). Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications, 38*, 14336-14345.
Karapetrovic, S., & Rosenbloom, E. S. (1999). A quality control approach to consistency paradoxes in AHP. *European Journal of Operational Research, 119*, 704-718.
Kou, G., Ergu, D., & Shang, J. (2014). Enhancing data consistency in decision matrix: Adapting Hadamard model to mitigate judgment contradiction. *European Journal of Operational Research, 236*, 261-271.
Lin, C., Kou, G., & Ergu, D. (2013). An improved statistical approach for consistency test in AHP. *Annals of Operations Research, 211*, 289-299.
Linares, P. (2009). Are inconsistent decisions better? An experiment with pairwise comparisons. *European Journal of Operational Research, 193*, 492-498.
Ma, L.-C., & Li, H.-L. (2011). Using Gower Plots and Decision Balls to rank alternatives involving inconsistent preferences. *Decision Support Systems, 51*, 712-719.
Saaty, T. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology, 15*, 234-281.
Saaty, T. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.
Saaty, T. (2003). Decision-making with the AHP: Why is the Principal Eigenvector necessary? *European Journal of Operational Research, 145*, 85-91.
Saaty, T. L. (1980). *Analytic Hierarchy Process*: McGraw Hill Higher Education.
Saaty, T. L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary. *European Journal Of Operational Research, 145*, 85-91.

Siraj, S., Mikhailov, L., & Keane, J. (2012). A heuristic method to rectify intransitive judgments in pairwise comparison matrices. *European Journal of Operational Research, 216*, 420-428.

Xu, Z., & Xia, M. (2014). Iterative algorithms for improving consistency of intuitionistic preference relations. *Journal of the Operational Research Society, 65*, 708–722.

Zhang, H., Sekhari, A., Ouzrout, Y., & Bouras, A. (2014). Optimal Inconsistency Repairing of Pairwise Comparison Matrices Using Integrated Linear Programming and Eigenvector Methods. *Mathematical Problems in Engineering, 2014*, 63-78.