Antoniades K, Ishizaka A. Experimental evaluation of the appropriateness of an interactive inconsistency correction, International Symposium of the Analytic Hierarchy Process 2014, Washington D.C., U.S.A.

EXPERIMENTAL EVALUATION OF THE EFFECTIVENESS OF AN INTERACTIVE INCONSISTENCY CORRECTION

Kyriacos Antoniades, Alessio Ishizaka
Portsmouth Business School
University of Portsmouth
Richmond Building, Portland Street,
Portsmouth, PO1 3DE
E-mail: kyriacos.antoniades@port.ac.uk, alessio.ishizaka@port.ac.uk

ABSTRACT

In a previous experimental study, it was observed that an automatic correction of inconsistency worsens the preference representation of the decision maker. In this paper, a new experimental study investigates if decision makers' preferences are better represented using an interactive inconsistency correction technique. The experimental results show that an interactive approach does not better represent decision makers' preferences for both the subjective and objective measures. Therefore, the interactive effort to reduce inconsistencies is not justified.

Keywords: AHP, Consistency, Goal Programming, Interactive, Preferences.

1. Introduction

The Analytic Hierarchy Process (AHP) is one of the most popular multi-criteria decision making (MCDM) method. Its success is widely accepted as it has been extensively implemented in many organizations to support their decisions. One of its main advantage stems from utilizing information redundancy to check for inconsistency. However, if inconsistency is too high, some of the pairwise comparisons may need to be reconsidered, before calculating priorities. Several techniques have been developed for this task. In this paper, the automatic correction of inconsistency is compared with an interactive correction. It is observed that neither the automatic nor the interactive approach do not better represent the decision makers' preferences, for both the subjective and objective measurable alternatives.

2. Hypotheses/Objectives

Linares (2009) found, in an experimental study, that automatic inconsistency correction with goal programming (Gonzalez-Pachon & Romero, 2004) worsen the ranking representation of decision makers. Successively, 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*." This paper attempts to answer the following research questions:

1. *Does the interactive approach improve consistency?*
2. *Does the interactive approach better represents participants' preference of rankings?*
3. *For the objective measures which preference of rankings is closest to the real measure?*

3. Research Design/Methodology

Two separate experiments are conducted in the following order:

a) *Decision with subjective alternatives*: Participants are asked to compare pairwise their preferred mode of transportation (Train, Coach, Taxi, Car sharing and Own car), to travel from Portsmouth to Gatwick airport to take a flight for a short week-end break.
b) *Decision with objective measurable alternatives*: Participants are asked to evaluate the closest distance of five cities (Cardiff, London, Edinburgh, Southampton, and Liverpool) to Portsmouth.

During the experiment, three rankings are produced by the participants:

1. Original Ranking (R_O), where the priorities are calculated by the eigenvector method without any inconsistency correction.
2. Automatic Ranking (R_A), where inconsistencies are corrected automatically using the goal programming method.
3. Interactive Ranking (R_I), where the decision analyst collaborates with the participant to detect and correct the most inconsistent pairwise comparisons to the least one.

The procedure is exactly the same for both experiments:

1. The decision problem is explained to participants.
2. The participant compares pairwise the five alternatives.
3. The Consistency Ratio (CR), the Original Ranking (R_O) and the Automatic Ranking (R_A) are calculated.
4. If the consistency ratio is acceptable, i.e. below 10%, the experiment terminates.
5. If the matrix is inconsistent, the inconsistency of each pairwise comparison, I_{ij} is calculated with:

$$I_{ij} = a_{ij} - \frac{p_i}{p_j} \tag{1}$$

   where, a_{ij} is the pairwise comparison of alternative i with j and p_i, p_j are their respective priorities (Saaty, 2003).
6. The participant is asked if (s)he wants to revise the most inconsistent comparison. If (s)he does not want to revise, then (s)he is asked if (s)he wants to revise the next most inconsistent comparison. If (s)he revises the comparison, the process restarts from point 4 until the inconsistency falls below 10% or the participant does not want to revise any of their entries.
7. The final Interactive Ranking (R_I) is calculated.
8. The participant is asked which of the three ranking (without knowing how they have been calculated) represent the best his/her preferences. At the end of the experiment, a questionnaire is completed with participant's demography details as well as feed-backs on the research study.

The program has been developed in Excel; based on a modified Excel template by Goepel (2013) with an embedded link to LINGO to calculate the Automatic Ranking (R_A) using the goal programming method.

4. Data/Model Analysis

A number of observations are made on 31 participants, these include:

C: Number of detections and corrections (Revisions) of the inconsistent pairwise comparisons.
N: Number of declined revisions. When participants do not wish to change their pairwise comparison; the next most inconsistent element is considered
CR_i: Initial Consistency Ratio (CR) at the first completion of the matrix.
CR_f: Final Consistency Ratio (CR), i.e. when it reaches an inconsistency below 10% or when the participant does not want to revise any of his/her entries.
ΔCR: Difference between initial and final Consistency Ratio (i.e. $\Delta CR = CRi - CR_f$).
PR: Participant's Choice of ranking (i.e. Original, Automatic or Interactive).
RR_A: Rank Reversal in the Automatic Ranking with respect to the Original Ranking.
RR_I: Rank Reversal in the Interactive Ranking with respect to the Original Ranking.
Close: Ranking which is close to the real measure.

Table 1. Participant's choice of rankings for subjective and objective measures.

Subjective Measures									**Objective Measures**									
#	**C**	**N**	**CR_i**	**CR_f**	**ΔCR**	**RR_A**	**RR_I**	**PR**	**#**	**C**	**N**	**CR_i**	**CR_f**	**ΔCR**	**RR_A**	**RR_I**	**Close**	**PR**
P5	5	2	29.71	7.36	22.35	yes	yes		P1	5	10	24.99	12.79	12.2	no	no	Interactive	
P6	2	0	19.54	8.62	10.92	no	no		P3	1	0	10.83	9.24	1.59	yes	no	Original	
P7	5	0	19.70	5.26	14.44	yes	yes		P10	1	0	11.58	8.70	2.88	yes	no	Original	
P9	4	0	27.43	9.99	17.44	no	no		P11	5	0	16.41	9.44	6.97	yes	no	Original	
P10	2	0	23.85	4.15	19.70	no	no		P12	1	0	11.04	8.03	3.01	yes	no	Original	
P14	2	0	10.74	8.20	2.54	yes	no		P15	1	0	14.47	8.67	5.80	yes	no	Original	
P15	7	3	16.09	9.49	6.60	yes	yes		P17	1	1	12.77	9.96	2.81	no	no	Interactive	
P19	7	4	85.94	9.79	76.15	yes	yes		P19	2	1	13.16	4.50	8.66	no	no	Interactive	
P20	5	10	25.84	20.81	5.03	no	yes	**Interactive**	P20	1	0	12.73	9.21	3.52	no	no	Interactive	**Interactive**
P13	2	0	21.56	9.81	11.75	yes	no		P25	1	0	11.84	8.97	2.87	yes	no	Original	
P16	3	0	15.56	6.84	8.72	yes	no		P26	2	0	14.86	9.20	5.66	no	no	Interactive	
P17	1	0	11.61	8.40	3.21	yes	no		P28	3	0	12.20	8.74	3.46	no	no	Interactive	
P23	1	0	11.21	10.00	1.21	no	no		P31	1	0	11.72	8.61	3.11	no	no	Automatic	
P27	1	0	13.08	8.14	4.94	yes	no	**Automatic**	P27	0	0	5.95	5.95	0.00	yes	no	Automatic	**Aut.**
P28	2	0	10.14	9.24	0.90	no	no		P2	2	10	34.87	30.9	3.97	yes	no	Interactive	
P31	5	10	21.23	15.86	5.37	yes	yes		P4	0	0	4.37	4.37	0.00	yes	no	Original	
P1	0	10	13.73	13.73	0.00	no	no		P5	0	0	3.45	3.45	0.00	no	no	Original	
P2	0	10	15.28	15.28	0.00	no	no		P6	0	0	7.39	7.39	0.00	yes	no	Original	
P3	1	10	30.97	28.30	2.67	no	no		P7	0	0	8.77	8.77	0.00	no	no	Original	
P4	7	2	20.63	9.53	11.10	no	no		P8	0	0	7.56	7.56	0.00	yes	no	Original	
P8	1	0	11.76	7.96	3.80	yes	yes		P9	0	0	1.26	1.26	0.00	no	no	Original	
P11	0	10	33.91	33.91	0.00	no	no		P13	0	0	6.08	6.08	0.00	no	no	Original	
P12	8	1	27.98	9.09	18.89	no	no		P14	0	0	6.67	6.67	0.00	yes	no	Original	
P18	0	0	4.43	4.43	0.00	no	no		P16	0	0	8.27	8.27	0.00	yes	no	Original	
P21	0	10	20.55	20.55	0.00	yes	no	**Original**	P18	1	0	10.52	7.20	3.32	no	no	Interactive	
P22	0	0	5.78	5.78	0.00	yes	no		P21	0	10	10.29	10.29	0.00	no	no	Original	
P24	12	10	33.85	19.52	14.33	no	yes		P22	3	1	17.48	8.65	8.83	no	no	Original	
P25	0	0	2.04	2.04	0.00	yes	no		P23	0	0	7.33	7.33	0.00	yes	no	Original	
P26	0	0	8.63	8.63	0.00	no	no		P24	3	1	13.95	8.15	5.80	no	no	Original	
P29	1	0	11.01	8.03	2.98	yes	no		P29	0	0	6.81	6.81	0.00	yes	no	Original	
P30	0	0	9.68	9.68	0.00	yes	no		P30	3	1	12.39	9.16	3.23	yes	no	Interactive	

The results obtained are shown in Table **1**. The shaded rows in Table 1 show experiments, where no revisions (C = 0) are made.

Nine participants of the subjective problem and thirteen participants of the objective problem are excluded for the first and second research question because no consistency correction method was used. The following hypotheses are tested:

Hypothesis 1: The interactive method improves consistency.

For the problem with subjective alternatives, it appears that improvements in consistency (n = 22) are disproportionately over-represented compared to no improvements in consistency (n = 0). The Chi-square test, χ^2 (1, n = 22) = 22.00, significance threshold $p < .05$, confirms that the interactive method significantly improve consistency.

For the problem with objective measures, it appears that the improvements in consistency (n = 18) are disproportionately over-represented compared to no improvements in consistency (n = 0). The Chi-square test, χ^2 (1, n = 18) = 18.00, $p < .05$, confirms that the interactive method significantly improve consistency.

Hypothesis 2: The interactive method better represent participants' preference of ranking.

For the problem with subjective measures, it appears that the interactive method (n = 9) is proportionately represented with the automatic method (n = 7) and the original method (n = 6). The Chi-square test, χ^2 (2, n = 22) = 0.64, $p > .05$, does not confirm the hypothesis.

For the problem with objective measurable alternatives, it appears that the interactive ranking (n = 13) is proportionately represented with the original ranking (n = 5) The Chi-square test, χ^2 (1, n = 18) = 3.56, $p > .05$, does not confirm the hypothesis.

Hypothesis 3: The original ranking is closest to the real values

For the problem with objective measurable alternatives, it appears that the preference rankings that are closest to the real measures is the original ranking approach (n = 20). It is disproportionately over-represented compared with the interactive ranking approach (n = 9) and the automatic ranking approach (n = 2). The statistical result, χ^2 (2, n = 31) = 15.94, $p < .05$ confirms that the original value is the closest to the real value.

5. Limitations

In this experimental study, only two decision problems were used. Other problems need to be analyzed to confirm our observations.

6. Conclusions

The statistical results show that the interactive method does improves consistency for both the subjective and objective measures. However, it is observed that the interactive

approach does not better represent the decision makers' preferences, for both the subjective and objective measurable alternatives.

Moreover, the results also show, that participants' preference of ranking which is closest to the real measure is given by the original ranking approach, which indicate that the effort to reduce inconsistencies using the interactive approach is not justified.

7. Key References

Goepel, K. D. (2013). Implementing the analytic hierarchy process as a standard method for multi-criteria decision making in corporate enterprises-A mew AHP excel template with multiple inputs. Retrieved from http://bpmsg.com/new-ahp-excel-template-with-multiple-inputs/

Gonzalez-Pachon, J., & Romero, C. (2004). A method for dealing with inconsistencies in pairwise comparisons. *European Journal of Operational Research*, *158*(2), 351–361. Retrieved from http://ideas.repec.org/a/eee/ejores/v158y2004i2p351-361.html

Linares, P. (2009). Are inconsistent decisions better? An experiment with pairwise comparisons. *European Journal of Operational Research*, *193*(2), 492–498. doi:10.1016/j.ejor.2007.11.030

Saaty, T. L. (1980). *Analytic Hierarchy Process* (p. 287). McGraw Hill Higher Education. Retrieved from http://www.amazon.co.uk/Analytic-Hierarchy-Process-Thomas-Saaty/dp/0070543712

Saaty, T. L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary? *European Journal of Operational Research*, *145*(1), 85–91. doi: 10.1016/S0377-2217(02)00227-8