

AGGREGATING PAIR-WISE COMPARISONS GIVEN IN SCALES OF DIFFERENT DETAIL DEGREE

V.V.Tsyganok *

E-mail: vitaliy.tsyganok@gmail.com

O.V.Andriichuk

E-mail: andreychuck@ukr.net

National Academy of Sciences of Ukraine
The Institute for Information Recording
Kyiv, Ukraine

ABSTRACT

In this paper we suggest an original approach to conducting individual pair comparisons during group decision-making (including AHP/ANP-based decisions). Under this approach every expert is given an opportunity to use the scale, in which degree of detail (number of points) most adequately reflects this expert's competence in the issue under consideration. Before aggregation all separate expert estimates (judgments) are brought to a unified scale, and scales, in which these judgments were built, are assigned respective weights. A respective instrument for pair comparison conduction has been developed, and an experiment has been organized. The experiment statistically proves that as a result of suggested technology usage, there is an increase in degree of correspondence between estimates, input by an expert, and his/her own notions on examination objects.

Keywords: group decision making, expert judgments, pair-wise comparisons, scales of different detail degree.

1. Introduction

The practice of expert examination conduction (including AHP-based examinations) indicates that there are certain difficulties arising when verbal scales are used for expert examination. Expert/decision-maker is often offered to use only one scale for pair comparisons. Judging from experience, in order to get thorough and undistorted data from an expert, (s)he must be offered to input estimates in a scale, which most adequately corresponds to his/her competence (awareness) level in the issue under consideration. The suggested research resolves the issue of using verbal scales with different degree of detail for each particular pair comparison, in order to ensure maximal credibility of knowledge obtained from an expert (expert information must be thorough and undistorted).

To ensure thoroughness of information obtained from an expert, we suggest using verbal scales of sufficient degree of detail: the more points the scale includes, the more information an expert can, potentially, input into a DSS, using this particular scale. To avoid information distortion (if an expert is unsure of the degree of dominance between objects in a pair, i.e. (s)he is not competent enough), we suggest giving experts the

opportunity to use scales with low degree of detail, or, even, refuse to estimate preferences in certain object pairs. Besides, in our research we also consider an important factor, influencing the level of expert information distortion – quantitative equivalent, corresponding to this or that point of a verbal scale. Correspondence between preference value input by an expert and this expert's notions about ratio of object weights on a pair is an issue of great importance, as it influences the credibility of expert data-based decision-making recommendations.

2. Literature Review

The key studies in the described area include the recent research by (Elliott, 2010), addressing the influence of a chosen quantitative scale upon correspondence between estimation results and expert's own notions. It was demonstrated, that scale selection has considerable impact upon the resulting decision variant estimate. Three quantitative scale types were analyzed, whose point values were assigned to fundamental scale points (Saaty, 2006) of two kinds, i.e. scales with 5 and 9 grades. Quantitative scales under consideration included integer, balanced (Salo & Hamalainen, 1997) and power (Stevens, 1957) scales. In contrast to research, described in the listed publications, in our study we suggest choosing a different scale for each single pair comparison, and not for all pair comparisons.

3. Hypotheses/Objectives

The purpose of this study is to prove, that to ensure obtaining of thorough and undistorted expert information on relation between objects (on estimates provided during pair comparisons), an expert should be given an opportunity to use scales with different degree of detail (accuracy). This hypothesis is based on a presumption that in every issue under consideration (and in every pair comparison) an expert has a different level of knowledge/competency/awareness. Each expert's competence level can correspond to a respective estimation scale: the higher the expert's competence is, the more detailed scale (s)he can use to adequately present his/her knowledge. According to the same principle, an uninformed/incompetent expert should have an opportunity to use a scale with small number of grades (including ordinal scale with only two values – “more” or “less”) for pair comparisons, or even refuse to compare objects in a pair because of incompetence. It is understandable that an expert judgment provided in a more detailed scale should be considered more significant than that same judgment provided in a less detailed scale, because in the first case the expert is more confident, and his self-estimated competence in the issue under consideration is higher. Consequently, if during pair comparisons an expert considers objects equal, this judgment can be considered the same as refusal to conduct this particular comparison (inability to evaluate preference of objects in a pair due to doubts/low competency in the issue under consideration). As we see, in verbal scales there is no real need for a grade “equal”/“no preference”, because if an expert chooses this value, (s)he might as well “skip” (refuse to estimate) respective preference. Anyway, the choice of “equal” preference value does not introduce any additional information on relation between objects.

Proof (confirmation) of any hypothesis in a weakly structured domain (in which we are conducting our research) is problematic, as there are absolutely no benchmarks to compare results with. That is why, the only way to confirm the hypothesis is an experiment using estimates provided by experts. Such an experiment is described in section 5 of this paper.

4. Research Design/Methodology

During the research a methodology and respective software tools were developed to conduct expert estimation based on the abovementioned approach. In group estimation every expert is offered to provide pair comparisons in verbal scales with different degrees of detail. Each particular pair comparison starts with the scale including only two values («Less» and «More») with an opportunity to refuse to provide the judgment – «No idea» (Figure 1 a).

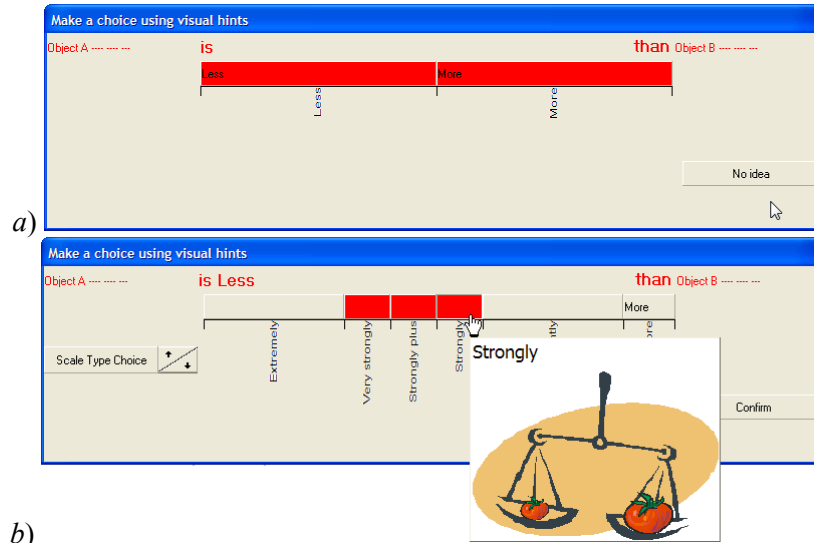


Figure 1 Software tools for gradual estimate precision increase

If ordinal comparison is provided (one of the values «Less» or «More» is selected) the expert is offered to gradually make the estimate more precise, and stop estimation at any stage («Confirm» button on Figure 1 b). In the process of this iterative procedure the final estimate is conducted in the scale, which most adequately corresponds to expert's competence in the issue of defining the preference relation between two particular objects. The final estimate may be provided in a scale including 2 to 8 grades.

It should be noted that the developed tool allows an expert to be sure that the quantitative equivalent really corresponds to this or that verbal phrase from estimation scale. Such confidence is achieved through providing user (expert) with interactive graphic tips (hints), allowing him to imagine the approximate relation between objects, and, thus, improve the degree of correspondence between the expert's personal notions and the information (s)he inputs during pair comparisons.

For aggregation of incomplete comparison matrices provided by a group of experts, when different comparisons can be conducted in scales with different accuracy, we suggest using the method known as enumerating all spanning trees with further averaging of priority vectors, calculated based on every tree (Tsyganok, 2010). Before calculation of priority vectors, all pair comparison matrix elements (judgments) are brought to a unified (most detailed) scale. During this process weights of particular judgments (pair comparisons) are taken into consideration. The weights depend on the degree of detail of scales the comparisons were provided in. The scale's degree of detail (informativeness) is

calculated according to Hartley's formula for quantity of information: $I = \log_2 N$, where N – number of points in expert estimation scale.

5. Data/Model Analysis

To confirm the hypothesis set forth in section 3 of this paper, an experimental research was conducted with real experts involved. As there are no benchmarks (model values), only the result of individual (not group) expert examination was analyzed. Every experiment participant was offered to select a subject domain, (s)he is competent in, and freely formulate an understandable goal. After that the participant (expert) was offered to formulate 5 to 7 factors making positive impacts upon the formulated goal. It should be noted that, since every expert chooses the subject domain (s)he is familiar with, (s)he must also be aware of contributions of each formulated factor into the goal's achievement.

During the next stage pair comparisons of impacts of formulated factors were conducted. Experts were offered to conduct further comparisons in 3 ways: in the fundamental scale with 5 grades ('Equivalent' (1), 'Moderately' (3), 'Strongly' (5), 'Very strongly' (7), 'Extremely' (9)), in the fundamental scale with nine grades ('Equivalent' (1), 'Weakly or slightly' (2), 'Moderately' (3), 'Moderately plus' (4), 'Strongly' (5), 'Strongly plus' (6), 'Very strongly' (7), 'Very, very strongly' (8), 'Extremely' (9)), and using the suggested tool. In order to minimize the correlation between repeated comparisons of same pairs of objects provided in different ways (every pair was compared three times – each time in a different way) the sequence of pairs presented to an expert for comparison was randomized.

After all pair comparisons were performed (3 pair comparison matrices were filled), 3 priority vectors were calculated. Eigenvector method was used to process matrices, built using the first two approaches, while to define a priority vector based on a matrix including comparisons provided in different scales, the so-called combinatorial (or spanning tree enumeration) method (Tsyganok, 2010) was used (particularly, its modification allowing for usage of different weights for different estimation scales).

At the final stage every experiment participant was offered to rank 3 priority vectors calculated for the factors (s)he formulated. Vectors were displayed as unsigned bar charts in random order. The participant was offered to rank the vectors according to their correspondence to his/her perceptions of quantitative relations between impacts of the formulated factors.

Result obtained by an expert (experiment participant) qualified only if pair comparison matrices satisfied sufficient consistency condition (C.I. value). Statistically credible results were obtained. These results are presented in Table 1. As a result of the research, we can conclude that in most of the analyzed cases, expert estimates obtained using the suggested technology, are more consistent with experts' individual perceptions of examination subject, in comparison to estimates, based on traditional estimation techniques (where fixed number of verbal scale grades is used). Consequently, wide implementation of the suggested pair comparison instrument in decision support technologies (including those using AHP/ANP) seems adequate.

Table 1 Comparative experimental research results

| Name of pair comparison technology | Number of participants, who assigned the specified rank to respective technology | | |
|--|--|-----|-----|
| | „1” | „2” | „3” |
| Fundamental preference scale with 5 grades | 10 | 15 | 37 |
| Fundamental preference scale with 9 grades | 12 | 33 | 17 |
| Technology suggested in the paper | 41 | 15 | 7 |

6. Limitations

Usage of the suggested tool for pair comparisons may require longer time during expert estimation, and, as a result, more resources, than traditional methods. This may result from the fact that more actions are required from experts during each pair comparison. But in reward we get higher credibility of expert estimates and recommendations to decision makers.

7. Conclusions

As a result of the research, we have suggested an expert estimation mechanism, allowing experts to use scales of different accuracy for each pair comparison. Relevance of the suggested approach is experimentally proven. It has been demonstrated that usage of the respective tool for pair comparisons allows us to improve the degree of correspondence between expert’s estimates and his notions of examination subject. This improvement results from the fact that experts use scales, whose accuracy is most consistent with their competency in every issue under consideration.

Implementation of the suggested expert estimation technology in combination with pair comparison matrix aggregation methods (including group methods) improves the credibility of AHP/ANP-based recommendations, given to decision makers.

8. Key References

Elliott, M.A. (2010). Selecting numerical scales for pairwise comparisons. *Reliability Engineering and System Safety*, 95, 750–763.

Saaty, T.L. (2006). *Fundamentals of decision making and priority theory with the analytic hierarchy process*. Pittsburgh, PA: RWS Publications.

Salo, A.A., & Hamalainen, R.P. (1997). On the measurement of preferences in the analytic hierarchy process. *Journal of Multi-criteria Decision Analysis*, 6, 309–319.

Stevens, S.S. (1957). On the psychophysical law. *Psychology Review*, 64, 153–181.

Tsyganok, V.V. (2010). Investigation of the aggregation effectiveness of expert estimates obtained by the pairwise comparison method. *Mathematical and Computer Modelling*, 52(3-4), 538–544.