Multidiszciplináris Tudományok, 15. kötet, 3. sz. (2025) pp. 11–19. ***https://doi.org/10.35925/j.multi.2025.3.2***

EXPLORING OPINION PATTERNS: BEYOND SIMPLE QUESTIONNAIRES

László Berényi
associate professor, University of Miskolc, Institute of Management Science
3515 Miskolc-Egyetemváros, e-mail: laszlo.berenyi@uni-miskolc.hu

Abstract
Due to the great variety of stakeholders and the rapidly changing environment, the multi-criteria decision problems are common in the field of management. Such problems require specific solutions due to their complexity and limited information. In an increasingly dynamic environment, managers should include tools and methods that can be flexibly adapted to address novel challenges. There are technical and personal barriers in case of a long list of items to set in order. A simple ranking may be impracticable, but pairwise comparison or the Q-sort method offers an old but rediscovered management tool for ranking and building opinion patterns. The study gives an overview of the opportunities.

Keywords: *survey, ranking, pairwise comparison, Guilford, Q-method*

1. Introduction

Regardless of working alone or in a team, following some general and industrial standards, applying methods forced by the supply chain, or failing these, managers are responsible for problem-solving. The wide range of challenges requires management to have a flexible toolbox with complementary elements that can be used to select the right solution to the problems encountered.

The problems facing an organisation can be well-structured or poorly structured. Well-structured problems assume knowledge of the situation, influencing factors, and possible outcomes. Such decisions are programmed and predictable. A planned increase in the volume of production allows us to predict the raw materials, workforce, or time needed for the new situation. At the same time, predicting the market needs and sales opportunities is a more complex challenge. The changes in customer preferences, regulation system, and competitor promotion are additional factors to consider, while limited information about the outputs and circumstances is available. Many situations require assumptions and estimates in order to identify the problem and the possible solutions:

- in the case of a choice between several options, their order of importance must be determined,
- involving several people relevant to the issue under consideration in the preparation and evaluation process also increases the thoroughness and acceptance of the decision,
- an appropriate method must be found for aggregating the results and weighting the factors evaluated, which necessarily entails a loss of information.

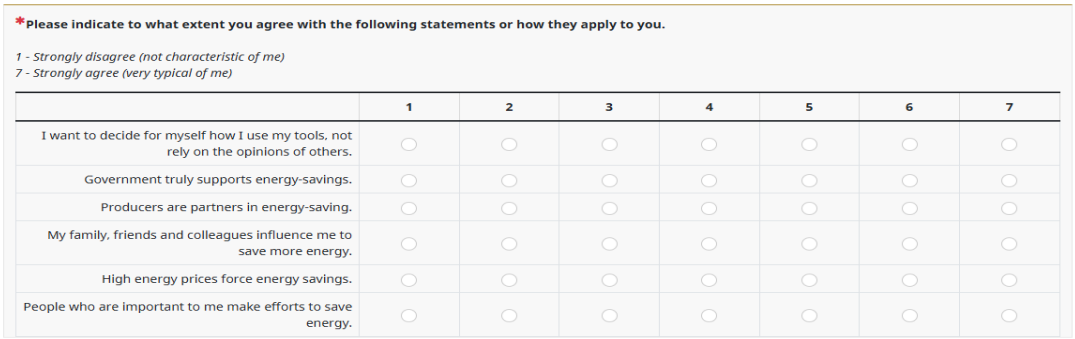
From a methodological point of view, the task quantifies characteristics and attributes that cannot be measured by engineering or economic indicators, like preferences, satisfaction, attitudes, and opinions. That is a complicated task when several perspectives and several factors must be taken into account together. Typically, a company sells its products (services) to a large number of consumers whose needs

are met to a greater or less extent by the features and functions of the products. The question can be formulated as follows: What features and characteristics should the product contain to enable a company to reach the consumers? In other words, what factors should be emphasized to achieve a given objective? To answer this question, it is crucial to know the value judgments of the stakeholders, particularly customers and users, and to know the average and extreme opinions. Organisational development actions face similar issues: dealing with risk assessment is complicated because actions will have different and competing impacts on the stakeholders.

The study offers a methodological contribution by presenting the limitations of some apparent ways of assessment and introducing relative ranking methods. Although these approaches are not novel, computerisation nowadays allows a simple and fast application, which increases their popularity. A wide range of free software is available for related data collection and calculations. However, attention should be drawn to the opportunities and their proper use since misinterpreting the methods may be harmful.

2. Limitations of simple scales

Business decision-making usually requires understanding mass phenomena, i.e., means and patterns of opinions, behaviours, or judgments. A simple scale, for example, scoring a statement between 1 and 5 (strongly disagree and strongly agree) seems to be a great way. A scale with five items is familiar to everyone from school grading and calculating some statistical indicators is easy. Of course, scale evaluation is multifaceted (see Babbie, 2020), using an even or odd number of options, mirrored questions, or individual endpoints *(Figure 1)*. Exploring opinion patterns is not impossible with a "simple" scale assessment, but numerous restrictions and misunderstandings must be taken into account.



***Please indicate to what extent you agree with the following statements or how they apply to you.**

1 - Strongly disagree (not characteristic of me)
7 - Strongly agree (very typical of me)

	1	2	3	4	5	6	7
I want to decide for myself how I use my tools, not rely on the opinions of others.	○	○	○	○	○	○	○
Government truly supports energy-savings.	○	○	○	○	○	○	○
Producers are partners in energy-saving.	○	○	○	○	○	○	○
My family, friends and colleagues influence me to save more energy.	○	○	○	○	○	○	○
High energy prices force energy savings.	○	○	○	○	○	○	○
People who are important to me make efforts to save energy.	○	○	○	○	○	○	○

Figure 1. *Scale assessment examples* (own edition)

By calculating the distribution, mean, standard deviation, kurtosis, and skewness, then applying variance analysis, factor analysis, or clustering, the diagnosis is ready. Comparison is also allowed: if respondents evaluate several products (attributes) using the same approach, the scores can be used to select better and worse solutions. There are several practical benefits of using scale assessments. Indeed, comparisons based on them systematically make the same mistakes for each measurement so that the results are ultimately comparable, but the quality manager should be aware of the limitations (Lengyelné, 2001; Gronhaug & Ghauri, 2011; Groves et al., 2009; Babbie, 2020):

– Scale assessment mostly uses an ordinal scale with no meaningful distances between each response. When the responses are later coded to numerical values, the distances between the numbers tend to be interpreted (as a ratio scale). What is the difference between "strongly disagree" and "tend to disagree", and is it the same as between "tend to agree" and "strongly agree"?

- As a consequence of the previous point, the results cannot (or rather should not) be used to calculate a mean, and especially not a standard deviation. However, in practice, we do so. That is a generally accepted practice in the social sciences, but we should be aware that it is only intended to present essential experience. The frequency of responses provides objective information.
- The middle value of the scale can be problematic if the questionnaire does not clearly define the middle value, the respondents' genuine opinion. If the term "neutral" is included, it is, in many cases, simply uninterpretable. There is also the term "do not know", which is convenient for the respondent but makes things difficult for the interviewer. On a five-point scale, if the survey shows a mean score of 3, it could mean that all respondents are moderately satisfied but also that just half are perfectly satisfied and half are perfectly dissatisfied.
- The five-point scale is a convention but not a requirement. A three-point scale may be more appropriate when asking young pupils how satisfied they are with something. An adult can distinguish between 5 and 10 options, but no more. Such a nuanced view is an excessive expectation.
- The problem of the middle value can be remedied by using an even-numbered rating scale, which will, in any case, lead the respondent to make a value judgment. The reason for using odd-numbered scales is familiarity. For example, one of our submitted articles was rejected because we used a six-point scale, and the editor thought that the averages calculated from this might confuse readers: a score of 3.00 on the five-point scale is just about the middle value, but on the six-point scale it is below. It is 'worth' to assume that readers expect the usual five-point scale.
- For example, instead of 1 and 5, the endpoints could be -2 and +2, but this does not change the characteristics of the distribution.
- The lack of knowledge or biases due to social expectations can make the results unreliable. If we ask respondents how important they consider certain global issues to be, we will find few low ratings. Filtering for biases requires extra effort, wasting the time of the researchers and the respondents.
- It should also be noted that not all scales are Likert scales that seem to be (see Babbie, 2020). The essence of Likert's method is the scale construction, not the grades.

Among the assessment scales, the semantic differential scale [1] is also worth highlighting for its practical utility and applicability *(Figure 2)*.

*The questions below ask about your preferred learning situation and style. Please indicate, in turn, which of the two characteristics you prefer for effective learning.

	much rather A	rather A	equal	rather B	much rather B	
A. learn in silence	○	○	○	○	○	B. recite the material
A. doing one thing at a time	○	○	○	○	○	B. doing several things at the same time
A. in class	○	○	○	○	○	B. independently
A. live broadcast	○	○	○	○	○	B. pre-recorded video
A. classroom instruction	○	○	○	○	○	B. online instruction
A. learn sitting still	○	○	○	○	○	B. move around while learning
A. hear, listen	○	○	○	○	○	B. read, see
A. theoretical material	○	○	○	○	○	B. exercise, example

Figure 2. *Semantic differential scale* (own edition)

It is ready to manage different assessment criteria for a given topic by asking the respondent to choose between two opposing positions (marked A and B in *Figure 2*) for each criterion, line by line. That method forces the formation of an opinion more than a Likert scale, and it can be used to profile preferences or satisfaction about a topic. It is useful for product design tasks and evaluating service attributes among experts or customers. Therefore, building the set of the items and the wording of the statements require special attention and effort. The opposite statements per line must be prepared clearly, as the evaluation criterion behind them is usually not shared with the respondents. Of course, the assessments per line are independent of each other; for example, "rather A" for the first item does not influence in any way the "rather B" judgment of the second item. The semantic differential scale provides an acceptable method for exploring opinion patterns, with additional steps of profiling sub-samples.

3. Ranking and pairwise comparison

The main limitations of scaled assessment can be overcome by asking the respondents to rank the items according to some criterion (importance, satisfaction, or else). A simple ranking ("put the items in order") is a good solution for a limited number of items: ordering 20-30 items may be exhausting and inaccurate.

In the case of a long list of items, it is advisable to look for another solution. In addition, the ranking scale does not provide information on the distance of each item at the level of the individual assessor, which may be important information for decision-makers. It must be noted that this information is also missed by using a scale evaluation, just as the coding system suggests. The total sample of the responses allows as follows:

- a rank correlation coefficient can be calculated, which provides information on the similarity or dissimilarity of individual ratings and
- the mean and distribution of the ranking scores can be used to assign weights, but more sophisticated procedures exist.

A further possibility for ranking is the use of pairwise comparisons (Kindler & Papp, 1977; Ramík, 2020; Mazurek, 2023), where the evaluator has to choose between only two items in a single step. This simplifies the execution of the task in exchange for having to perform several evaluation tasks: for n items, a total of $(n \cdot (n-1))/2$. In other words, a pairwise comparison of 3 items requires 3 pairwise evaluations, 4 items require 6 pairs, 5 items require 10 pairs, and 15 items require 105 pairs.

The inconsistency – an ambiguous order of preference – is still present in pairwise comparisons but is not hidden compared to simple ranking. For example, if one likes bananas more than apples and apples than pears, it means one likes bananas better than pears. If this order of preference is not violated in the case of the banana-apple-pear triad, it is called a consistent case. In practice, however, it cannot be ruled out that one prefers the latter, choosing between banana and pear. That is the inconsistency that complicates the calculations, but it is a natural phenomenon. In the case of a simple ranking, the evaluator gives a ranking and, at most, if the way of asking the question allows it, notes the specificity. In the case of pairwise comparisons, the presence of inconsistency can be detected afterward.

4. Guilford weighting for pairwise comparisons

Guilford's procedure (Guilford, 1936), dating from the 1930s, is a scale transformation solution that assigns weights to items on an interval scale in conjunction with pairwise comparisons. A powerful assessment tool that can be used for product, process, or organisation development issues. Although the

method needs a special questionnaire prepared for pairwise comparison, the data collection can be automated through an online questionnaire. Then, a template can be prepared in any spreadsheet software *(Table 1)* to calculate the data and show the results automatically. The main steps of the procedure are (based on Kindler and Papp, 1977):

- The definition of the research question and items to be assessed will be followed by the preparation of the data collection questionnaire. Ross's (1934) optimal arrangement of the order of the items is widely used to avoid an item being "too often" close to each other in the questionnaire.
- Collecting individual ratings, compiling the preference matrices, and calculating the consistency indices. In the evaluation, a maximum of $(n^3–n)/24$ decision triples for an odd number n of items and $(n^3–4n)/24$ for an even number of items. The individual consistency score is the ratio of the number of triples (d) that actually exist in the evaluation to the maximum number of triples $(1–(24d/(n^3–n)$ for odd items and $1–(24d/(n^3–4n)$ for even items. The result of the consistency test is used to decide whether to consider the responses of a particular evaluator in further analyses.
- The individual evaluations with the appropriate consistency level are aggregated, and the aggregate preference matrix is constructed.
- Determine the preference ratios from the preference frequencies (a) and transform them into u values of the standardized normal distribution. Statistical software or tables describing the distribution (see, for example, Kindler & Papp, 1977; Babbie, 2020) can be used to assist in the transformation.
- Express the values of the standardized normal distribution u in % (Z). This value is the weight on the interval scale of the element.

Table 1

Preference matrix and calculations, example (own edition)

	hydro-power	solar energy	nuclear energy	wind power	biomass energy	*a*	*(a+m)/2*	*P*	*u*	*Z*
hydro-power	**–**	**9**	**48**	**31**	**51**	*139*	*193.5*	*0.355*	*–0.37*	*0*
solar energy	**100**	**–**	**71**	**96**	**86**	*353*	*407.5*	*0.748*	*0.67*	*100*
nuclear energy	**61**	**38**	**–**	**57**	**59**	*215*	*269.5*	*0.494*	*–0.01*	*34.4*
wind power	**78**	**13**	**52**	**–**	**65**	*208*	*262.5*	*0.482*	*–0.05*	*31.3*
biomass energy	**58**	**23**	**50**	**44**	**–**	*175*	*229.5*	*0.421*	*–0.19*	*16.6*
sum	*297*	*83*	*221*	*228*	*261*	*1,090*				

A careful interpretation of the results is important. However, an exact value is assigned to the items, measured on an interval scale. That means the least preferred item always bears 0, and the most preferred is 100. The positions of the other items give valuable information, but ultimate decisions or comparisons between different samples are not allowed. For a ratio-scale result, the AHP method (Saaty et al., 2022) can be suggested (not processed in that paper).

5. Q-sort method

The Q-sort method (Q-methodology, Q-sort ordering) was developed in the 1930s (Stephenson, 1935; Stephenson; 1953) based on the problems of the traditional (still used today) factor analysis method. The traditional approach examines mass phenomena, works with many respondents, and only gives good results under a large number of statistical conditions. However, the conditions are often not met, and from a qualitative point of view, examining individual characteristics is crucial in identifying the causes of problems.

The Q-sort method is suitable for exploring opinion samples. It does not require a large sample, representative data collection, and a normal distribution of responses, i.e., it can "bypass" all the conditions required for traditional factor analysis and clustering. According to Brown (1996), the power of the method is based on the recognition that there are a limited number of basic patterns of opinion about a given topic so that even by asking a few people, we can learn about a large proportion of the patterns. Suppose that there are two people with precisely opposite opinions. If we ask a third person, their opinion is either the same as one of the previous two or somewhere in between. The following person involved has an opinion close to one of the three, and so on. Involving experts in the ranking makes it possible to find the dominant opinions quickly.

The method was initially used introspectively to explore psychological factors but later became popular for systematising the opinions of several individuals. The method is computationally intensive – which is why it was less widespread in the social sciences and economics – but is enjoying a renaissance in the 21st century. Several free software are ready to support data collection and analysis. A comprehensive solution is offered, for example, by Banasick's (2023) procedure and software.

The Q-sort method starts with formulating a problem or question based on which the respondents are to rank different factors (statements, characteristics). The literature suggests a range of 20 to 100 items to be ranked. Although the method handles inconsistent responses robustly and can handle the task mathematically, scoring multiple items can be time-consuming and tedious for respondents.

Data collection requires a specific questionnaire or template *(Figure 3)* that clearly outlines the respondent's relative opinion of each statement in relation to all other statements, presenting a holistic order with integrated trade-offs (Zabala & Pascual, 2016), i.e., it provides the possibility to evaluate certain factors at the same level. Data can be recorded on paper or online. The advantage of the latter solution is that data aggregation can be automated, data entry errors can be avoided, and calculations can be accelerated.

Respondents check cards describing the attributes or items they receive for the assessment, which they have to place in the template shown in *Figure 3*: on the right, those they consider more important in line with the question asked, and on the left, those they consider less important. The cards that are rated similarly are placed one below the other. Preliminary sorting can also be done by dividing the cards into three groups (e.g., less, uncertain, more). The pre-sorting does not necessarily mean that the cards must be placed in the template; respondents can change their minds. The main steps of the analysis:

- calculating correlations based on the initial data matrix of ratings,
- determining the number of factors (based on eigenvalues and the elbow method),
- calculation of (rotated) factor weights,
- calculation of rankings and statistical characteristics,
- analysis of common and discriminant factor statements,
- presentation of samples of opinions according to the final factors.

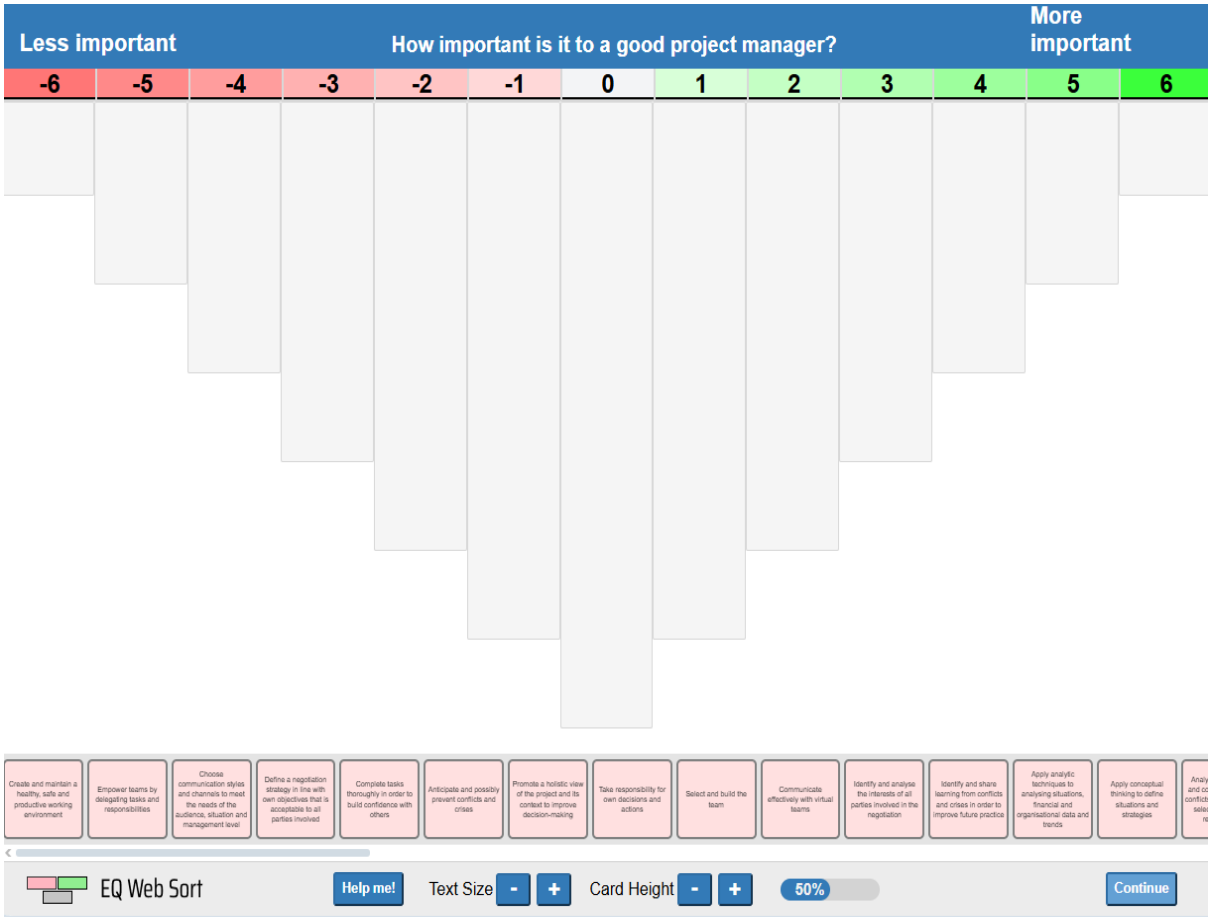


Figure 3. *Q-sort method "questionnaire" in EQ-Web sort template* (own edition by Banasick, 2023)

Although the Q-sort method uses statistical solutions, interpreting the results requires creativity. The calculations are guided, but the choice of the number of factors or the method of factorisation is difficult to algorithmize, as is the presentation of the practical utility of the resulting opinion patterns. Statistical details include ranking and Z-scores by items, correlations between the opinion patterns, and distinguishing items of the opinion patterns. The software provides a visual representation of the factors *(Figure 4)*. The results provide a quick overview of the most important or preferred factors (on the right side) or the least important ones (on the left side), in addition to which the significance levels and the consensus or distinguishing nature of the item can be shown compared to other factors in the model.

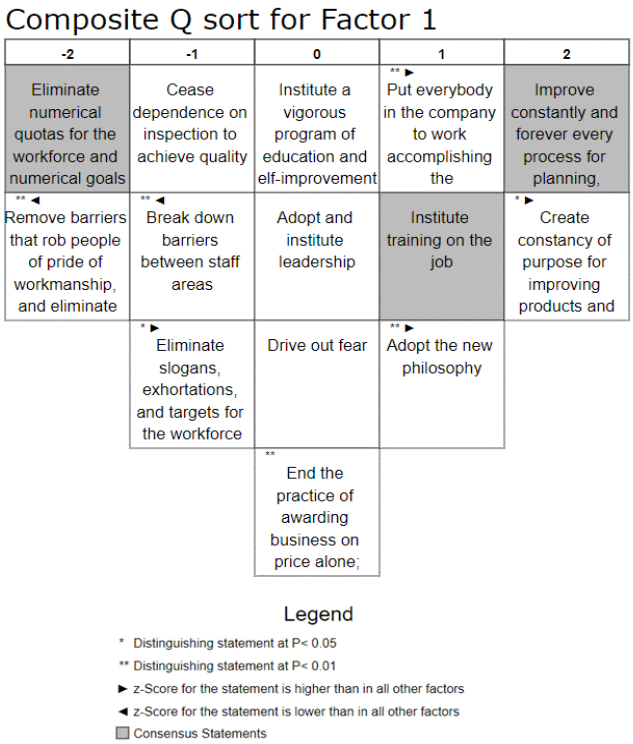


Figure 4. *Factor visualisation example* (Berényi, 2024)

Highlighting the patterns of opinion presented by the Q-sort method, particularly the common and distinctive statements, is important for developing appropriate responses. Similar features or statements across all groups can be used as a common point, while the distinctive items show aspects of developing product variants.

6. Summary

Ranking, evaluation, scales, and indices are common tools for decision-making support. It is worth going beyond simple evaluation tools to quantify complex problems with novel content. It is popular to quantify opinions using quick and easy scaling evaluation, but the reliability of the results is questionable. Ranking can produce "better" results, but the ability to weighting the evaluation factors is limited, and the presentation of the results is less impressive. There are various methods available for that purpose, usually typically with a grounded IT support for data collection and analysis. A comprehensive overview of the methods goes far beyond the scope of one paper. It is limited to demonstrating the opportunities of a pairwise comparison method, which is particularly suitable for ranking a few items, and the Q-methodology for broader purposes. The Guilford method based on pairwise comparisons with interval scale weights has been successfully used to support product development tasks. In contrast, the Q-sort method offers the possibility to identify typical opinion patterns for a large number of evaluation factors. In a rapidly and unpredictably changing world, implementing their approach enables managers to quickly and easily explore opinion patterns.

References

[1] Babbie, E. (2020). *The Practice of Social Research*. (15th edition). Cengage Learning.

[2] Banasick, S. (2023). Ken-Q Analysis (Version 2.0.1). **https://doi.org/10.5281/zenodo.8310377**

[3] Berényi, L. (2024). Assessing the actuality and relevant importance of Deming's 14 points. *Multidiszciplináris Tudományok: A Miskolci Egyetem Közleménye*, 14 (2), pp. 3–13. **https://doi.org/10.35925/j.multi.2024.2.1**

[4] Brown, S. R. (1996). Q methodology and qualitative research. *Qualitative Health Research*, 6 (4), pp. 61–567. **http://dx.doi.org/10.1177/104973239600600408**

[5] Gronhaug, K., Ghauri, P. (2011). *Kutatásmódszertan az üzleti tanulmányokban*. Budapest, Akadémiai Kiadó.

[6] Groves, R. M., Fowler F. J. Jr., Couper, M. P., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey Methodology*. Hoboken, Wiley.

[7] Guilford, J. P. (1936). *Psychometric Methods*. New York, McGraw-Hill Book.

[8] Kindler, J., Papp, O. (1977). *Komplex rendszerek vizsgálata. Összemérési módszerek*. Budapest, Műszaki Könyvkiadó.

[9] Lengyelné Molnár, T. & Tóvári, J. (2001). *Kutatásmódszertan*. Eger, Eszterházy Károly Főiskola Médiainformatika Intézet.

[10] Mazurek, J. (2023). *Advances in Pairwise Comparisons: Detection, Evaluation and Reduction of Inconsistency (Multiple Criteria Decision Making)*. Berlin, Springer. **https://doi.org/10.1007/978-3-031-23884-0**

[11] Ramík, J. (2020). *Pairwise Comparisons Method: Theory and Applications in Decision Making*. Springer, Berlin. **https://doi.org/10.1007/978-3-030-39891-0**

[12] Ross, R. T. (1934). Optimal orders for the presentation of pairs in the method of paired comparisons. *Journal of Experimental Psychology*, 23 (6), pp. 651–660. **https://doi.org/10.1037/h0058340**

[13] Saaty, T. L., Vargas, L. & St, C. (2022). *The Analytic Hierarchy Process*. Berlin, Springer .

[14] Stephenson, W. (1935). Correlating persons instead of tests. *Journal of Personality*, 4 (1), pp. 17–24. **https://doi.org/10.1111/j.1467-6494.1935.tb02022.x**

[15] Stephenson, W. (1953). *The study of behavior: Q-technique and its methodology*. University of Chicago Press.

[16] Zabala, A., Pascual, U. (2016). Bootstrapping Q Methodology to improve the understanding of human perspectives. *PloS ONE 2016*, 11, e0148087. **https://doi.org/10.1371/journal.pone.0148087**