Innovative Expert Methods in Strategic Decision Making

DOI: 10.12776/QIP.V26I1.1643

Dagmar Gombitová, Dušana Dokupilová

Received: 2021-01-14 Accepted: 2022-03-28 Published: 2022-03-31

ABSTRACT

Purpose: This study introduces managerial techniques applied for the first time in the high-level strategic public policy decision-making process in Slovakia with an aim to assess the strategic decision-making of groups of experts in a methodologically supported environment. It compares groups of internal analysts and external specialists and should demonstrate the extent to which these two groups are able to process problems analytically and suppress intuition.

Methodology/Approach: Multi-criteria decision methods are used when deciding on complex problems. One of the most popular and most frequently used is the Analytic Hierarchy Process. Application of this method enables measurement of preference consistency, and its relationship with cognitive reflection.

Findings: Consistency of judgement was very similar in both groups. The prioritisation of measures resulted in a similar set of priorities determined by both groups. The assumed relationship of consistency and cognitive reflection score and/or overconfidence was not detected, and decision makers proved to be well calibrated.

Research Limitation/Implication: The main limitation of our research was the small sample size of decision makers, which complied with the requirements of the decision method, but was not sufficient to confirm the statistical validity.

Originality/Value of paper: The introduction of the multi-criteria decision method into decision-making for public policy strategies combines practical policy exercises with scientific research on high-stakes decisions and enables to carry out participatory decision-making process with relevant stakeholders.

Category: Research paper

Keywords: managerial innovation; decision making; consistency; analytic hierarchy process

1 INTRODUCTION

Decision-making in public policy is complex and often costly, with a high degree of uncertainty and risk, as well as often having long-term consequences for a large number of people. Current public policies should ensure that this process is participatory, based on evidence and applies the latest scientific knowledge on decision-making. Strategic decision-making involves the selection of the best among the possible alternatives, although this process is largely associated with uncertainty resulting from inadequate knowledge and excessive complexity. The framework for high-stakes decision-making should therefore involve a methodology for choosing the best among the alternatives; this should take into consideration future situations, use techniques to assess alternatives, understand the behaviour of the system and incorporate expert opinions into the process (Bhushan and Rai, 2007). A review of the literature and common practice in Slovak high-stakes decision-making suggests that there is a research gap. As mentioned by Bačová (2010), there is no unified approach or guidelines, and not much is known about the strategic decision-making in Slovakia.

Our research introduced new managerial techniques – that is, a new way of organising the participatory decision-making process and assessing the strategic decision-making of groups of experts in a methodologically supported environment. Our study demonstrates the extent to which groups of experts/decision makers are able to process problems analytically and suppress intuition and/or to confirm to what extent this process can be considered rational in terms of economic theory. This was examined through consistency and its relationship with cognitive reflection and overconfidence. The contribution of our research lies in the examination of the real-life strategic decision-making of experts through a well-established scientific method, Analytic Hierarchy Process (AHP). AHP applications have grown exponentially in last decades (Emrouznejad and Marra, 2017; Zyoud and Fuchs-Hanusch, 2017), and it has been used in many policy areas, such as strategic socio-economic, technology or environmental decisions in numerous countries (Subramanian and Ramanathan, 2012).

There are a few documented cases of the use of AHP in the Slovak environment, but these cases did not concern public policy decisions (Peregrin and Karahuta, 2014; Šoltés and Gavurová, 2014). To the best knowledge of the authors, our real-life experiment was the first attempt in this geographic area to examine and compare the consistency of two expert groups in high-stakes decision-making. In the next part we briefly introduce the background of the experiment and a literature review of the main concepts applied in the research. The following methodological section introduces our approach focused on the goal of the experiment. The results section presents the calculations, and the concluding chapter summarises the main findings and states the limitations of the research.

The authors of the paper closely cooperated with the Ministry of Investment, Regional Development and Informatisation in Slovakia (MIRDI) and were

assisting with the setup of priorities for future support from European Union sources. The new programming period for the use of the European Structural and Investment Funds in Slovakia has been planned since 2019. The Partnership Agreement as an umbrella document setting future strategic priorities for support from the European sources in the new programming period 2021-2027 was being prepared. During its preparation, the pilot testing introduced the process of high-stakes decision-making utilising AHP for the first time (Baláž, Dokupilová and Filčák, 2021; Dokupilová et al., 2020). Two years later, this novel instrument for decision-making at a high strategic level was repeated. The implementation of one Operational Programme Slovakia 2021-2027 was agreed, which largely followed the structure set out in the Partnership Agreement; however, new measures were introduced to the political objective dealing with social issues, so new prioritisation measures was necessary.

1.1 Rational Decision-Making

New decision-making theories respect the fact that decisions are not strictly governed by the rules of mathematical logic, but that various heuristics/shortcuts are applied in the decision-making process, which manifest as deviations from rational thinking. Heuristics are commonly defined as cognitive shortcuts or general rules that simplify decisions, especially under conditions of uncertainty. Contemporary behavioural science knows several dozen heuristics (Kahneman, 2011), and they can lead to cognitive bias. Such distortions are sometimes appropriate because they are part of people's adaptive responses to situations (Lockton, 2012). Cognitive biases or deviations are repeated systematic deviations from the norm or errors that affect our judgment and decision-making. The terms heuristics and cognitive biases generally refer to the strategies used in decision-making; they differ from the rational model of decision-making. Rational decision-making is not seen as synonymous with intelligence, so it cannot be measured as IQ. Rational decision-making is inherently very complex and cannot be assessed through a single characteristic or variable. To assess cognitive abilities, the relationships of these abilities with heuristics and biases are correlated (Čavojová and Hanák, 2014; Toplak, West and Stanovich, 2014). Given the presumed multifaceted nature of rational decision-making, it is important to determine which measures are most appropriate for assessing rationality. One of the most frequently used measure is the Cognitive Reflection Test (CRT), which is a tool to assess cognitive abilities.

1.2 Decision-Making Methods

There are numerous methods applied in decision-making. The introduction of a participatory approach reduces the selection scope. As the most suitable approach for analysing major development challenges, the Delphi method was applied in the Slovak pilot project mentioned above. 'Classical Delphi' generates consensus on major development challenges, and 'political Delphi' encourages structured public dialogue to generate policy alternatives for further prioritisation

(Dokupilová et al., 2020). When decisions are made for complex problems, a multi-criterion decision-making (MCDM) can be applied. MCDM methods are designed to suppress intuition and 'objectify' the decision-making process as much as possible. MCDM aims to select the best alternatives from a set of many potential candidates or alternatives. The order of the alternatives is based on the evaluation of several conflicting criteria, which MCDM organises. The final number of alternatives is considered and evaluated according to the different criteria. MCDM has been applied in various fields (Wallenius et al., 2008), and about 100 MCDM methods have been used in business and the public domain (Danesh, Ryan and Abbasi, 2017; Velasquez and Hester, 2013). A meta-analysis of expert decision-making methods (Danesh, Ryan and Abbasi, 2017) indicated that one of the most popular and the most frequently used methods is AHP (Saaty, 2008), which was developed as a reaction to the lack of common, easily understood and easy-to-implement methodologies for complex decisions. AHP has found use in business, government, social studies and many other domains involving important and complex decisions in which choice and/or prioritisation is needed (Bhushan and Rai, 2007). AHP has been proved (Danesh, Ryan and Abbasi, 2017; Velasquez and Hester, 2013) to be a theoretically sound, tested and accepted tool.

This type of decision-making involves the participation of experts. Expert decision-making consists of the identification of alternatives and the selection of the most suitable ones based on the criteria and preferences of the decision makers. The expert decision-making method has several properties (Vidal, Marle and Bocquet, 2011) and should have several criteria for decision-making, including qualitative ones, and determine the order of priorities for alternatives based on the selected criteria. At the same time, the method should be reliable, accessible and understandable even to those decision makers who are experts in their field but do not have extensive experience in applying expert decision-making methods. Given the importance of the decision-making problem and the structure of the group of decision makers, when choosing a decision-making method, we consider the following factors: reliability, accessibility and the complexity of the decision-making alternatives. Performing pairwise comparisons of individual alternatives is to be preferred for all criteria, as it is easier and more accurate to express one's opinion on only two alternatives than simultaneously on all the alternatives (Ishizaka and Labib, 2011). For pairwise comparisons, AHP provides the possibility of creating one's own set of criteria that best capture the goal. This decision-making technique helps to alleviate any subjectivity or intuition involved in decision-making. In general, it is completely natural for people to follow intuition or to be subjective, and in some respects, it is a remarkable survival technique that can lead to quick decisions based on personal experience, but biased.

1.3 Expert's Decision-Making

In the pilot programme, decision-making had to be participatory (EC Directive request), and due to its nature, it required involvement of experts. The term 'expert' (also specialist and/or researcher) is very broadly defined as an individual/person who has some special knowledge and skills most often assessed based on criteria such as length of practice or education/training. Several research studies have concluded that experts can make the same mistakes in decision-making as lay people, they have the same limited cognitive capacity and use heuristics (Baláž, 2009; Bazerman and Moore, 2013; Kahneman, 2011; Larrick and Feiler, 2015; Tetlock, 2005). Further research has, however, presented a more positive view of expert decision-making and emphasised that it has room for improvement (Burgman et al., 2011; Hutton and Klein, 1990; Shanteau and Stewart, 1992; Wagenaar and Keren, 1985). Experts are usually unaware of subjective influences and often overestimate their abilities. Nevertheless, the use of experts in decision-making is important and useful for several reasons: they possess more appropriate decision-making skills, are faster and more efficient (Gilmour and Corner, 1998). Appropriately selected methods increase the accuracy and calibration of experts' judgement. Studies have confirmed that the selection of experts based on their expertise or experience is crucial, as their expertise declines dramatically outside their field. Some professions (e.g. chess players, meteorologists) are much more consistent and accurate, but they mostly study physical phenomena (as opposed to human behaviour, society) or recurring events where decisions are made intuitively and have a decision-making apparatus or tool (Shanteau and Stewart, 1992). Group decision-making is also beneficial, because these estimates are better than the estimates of individuals (Saaty and Peniwati, 2013). Our selection of experts was in line with the natural selection of experts when a participatory approach is required in Slovakia. The terms 'expert' and 'specialist' are often used as synonyms, but they should be distinguished. A specialist is a person who has the capacity to substantiate many suggestions in a subdomain or area. Specialists are therefore a subset of experts. Experts do not have to be specialists; an obvious example of this is a general practitioner in medicine who is an expert but not a specialist (Weinstein, 1993).

Two groups of experts participated in the experiment: permanent experts (analysts) and ad-hoc experts (specialists). The first group consisted of analysts from the analytical units of the relevant ministries and other state institutions. Analytical units are part of these institutions, and their role is to provide analytical services for the needs of the ministry or other central state administration body to which they belong. The addressed analytical units were from the Ministry of Labour, Social Affairs and Family; Ministry of Education, Science, Research, and Sport; and the Ministry of Health of the Slovak Republic. Staff of the Institute of Financial Policy and the Value for Money Department from the Ministry of Finance, and Analytical Unit of the Office of Government, who also deal with the sectoral strategies were also included. The employees of

the analytical units should thus have a very good overall overview of what is happening in the respective sectors. The second group of experts (specialists) was composed of knowledgeable and experienced people from the professional community. These specialists are more engaged in specific issues from the relevant sector. Unlike the first group of analysts, the second group has a much narrower professional focus.

2 METHODOLOGY

2.1 Process

The groups of decision makers and/or potential respondents were invited by the MIRDI to take part in the experiment. The aim of the Ministry was to set out the priority measures for the political objective Social Europe in the Operational Programme Slovakia 2021-2027. To enable the use of AHP to elicit policy priorities, an online assessment tool was developed. The outputs were recorded in Qualtrics XM. The initial part of the tool consisted of the decision-making exercise. The political objective covered three main topics of social sector: employment and labour market; education; and social inclusion and social-health services. The respondent could choose to prioritise measures in one or more topics. Each topic was divided into three or four measures (alternatives), which were prioritised. A detailed description of each measure comprising the rationale, target group, expected impact and type of supported activities was made available during the assessment process. The respondents had to read and process information and carry out pairwise comparisons for all alternatives. This process usually took from 30 to 90 minutes and generated a significant cognitive burden for the respondents. First, the standard weights of the criteria were set. The respondents started with mutual comparison of the three determined criteria. A pairwise comparison for each pair of criteria defined which is preferred, as well as by how much on a scale from 9 (criteria A is 9 times more important than criteria B) to 9 (criteria B is 9 times more important than criteria A). Subsequently, participants compared all alternatives according to the individual criteria. The maximum number of AHP criteria was limited to three and the number of alternatives to four as the overload (3+18 comparisons per topic with 4 alternatives, 3+30 comparisons with 5 alternatives) can reduce the quality of policy evaluation. For each criterion, a matrix of preferences was compiled for each alternative, which consisted of standardised weights of alternatives according to the individual criteria. Priorities were set by a rigorous mathematical process the priority vector of weights w. (Mu and Pereyra-Rojas, 2017). The vector method (Saaty, 2008) was used to calculate the priority vector in the AHP method.

When the decision-making process was completed, the respondents were asked to resolve an additional task. This consisted of 10 verbal cognitive reflection tasks (Sirota et al., 2021). Subsequently, the last task provided an estimated

number of correct answers in a cognitive reflection test. The online tool was accessible for one month and data collection took place in November 2021.

2.2 AHP Criteria

AHP breaks down the decision-making problem into elements and levels according to common characteristics. It creates a logical hierarchy that systematically evaluates pairs of alternatives according to specific criteria (Berrittella et al., 2011). To reflect the implementation structure of the Operational Programme, three criteria were defined for what the individual decision makers are assessing according to a given criterion:

1. Relevance determines the alternatives that could significantly affect the social, economic and environmental development of the Slovak Republic. The alternative (measure) with the highest priority should significantly help the country to face major societal challenges in the next decade. Failure to implement this alternative would lead to serious economic, social and environmental consequences and irreversible changes.
2. Urgency sets the alternatives that must be implemented immediately or at least in the shortest possible time, as this alternative may be a necessary condition for the introduction of other alternatives. Postponing this alternative to a later date would lead to serious economic, social and environmental consequences and irreversible changes.
3. Feasibility concerns the alternatives the Government of the Slovak Republic, individual ministries and their agencies can put into practice. No alternative alone is likely to change people's thinking, stop demographic change or solve the global climate crisis; however, some alternatives can have a significant impact on our society, economy and environment, although in the past, we have failed to implement them. We therefore need to know if it is possible to implement a given alternative with the current understanding of the issue.

2.3 Consistency

Many decision makers are inconsistent in their decision-making, especially when it comes to making comprehensive assessments of complex information. When asked to evaluate the same information twice, they often respond differently, and they are not consistent when choosing preferences (Tversky, 1969). It is evident and common that people systematically violate the principle of consistency even in relatively simple decisions (Dawes, 1979; Kahneman, 2011), but a certain level of inconsistency in decision-making is acceptable (Ishizaka and Labib, 2011). As the consistency of preferences is considered one of the basic elements of a rational decision-making process, the rational actor should be an actor with consistent preferences (Sičáková-Beblavá, 2015). As people violate this principle, we examined the consistency of experts in the process of comparing

alternatives. Consistent decisions meant the preference of the most advantageous alternative while applying transitivity, which is the basic principle of decision models based on pairwise comparison (Wu and Tu, 2020). The method transforms mostly empirical data, into numerical values, which are further compared (Vargas, 2010). More precisely, if one considers alternative A to be three times more important than B and B twice as important as C, then alternative A must be (3*2) six times more important than C. According to the principle of transitivity, individuals should have a well-defined preferential structure – that is, they should be consistent. Along with consistency, it is also possible to test other psychological characteristics.

2.4 Cognitive Reflection

The Cognitive Reflection Test (CRT) is a simple test and functions effectively as a demonstration of cognitive ability (Frederick, 2005). CRT is a good measure of incorrect information processing and measures the tendency to replace fast, intuitive, but incorrect choices (Oppenheimer and Thomson, 2016; Sirota et al., 2018; Sleboda and Sokolowska, 2017; Toplak, West and Stanovich, 2014). Toplak, West and Stanovich (2011) found that CRT estimates rational behaviour better than measures of cognitive ability, thinking dispositions, and executive functioning. The CRT test is also related to academic results (Frederick, 2005; Welsh, Burns and Delfabbro, 2013) and is therefore used to measure analytical cognitive style (Čavojová and Hanák, 2014). The original test consists of three simple numeric tasks that evoke intuitive, albeit incorrect, answers. Toplak, West and Stanovich (2014) presented its extended version with seven tasks. Sinayev and Peters (2015) provided evidence that numeracy skills are strongly associated with reflection ability, and this type of test may also cause a gender performance gap (Juanchich, Sirota and Bonnefon, 2020). New versions of the CRT have been developed (Primi et al., 2016; Sirota and Juanchich, 2018) that assume not all individuals are mathematically proficient. To avoid the problems with numeracy and the fact that the original test is well known, the new tested version of verbal CRT was used (Sirota et al., 2021).

2.5 Overconfidence

Cognitive reflection predicts several heuristics, and people with lower cognitive reflection are significantly more likely to be subject to overconfidence, the illusion of control and conservatism (Noori, 2016). Overconfidence is one of the most common heuristics and exists in several forms. The most frequently studied form is overestimation, which is a systematic overestimation of one's abilities compared to reality. Previous research has confirmed that less competent people tended to overestimate their abilities, while more competent people tended to underestimate them (Dunning et al., 2003; Kruger and Dunning, 1999; Pennycook et al., 2017). It can therefore be assumed that more competent experts would not be too confident.

2.6 Sample Size and Structure

There is a wide-ranging debate in AHP groups about group size, as a high number of experts may increase the inconsistency of judgements in the AHP. Saaty and Özdemir (2014) point to the consistency of experts and the validity of their experiences and recommend a maximum of seven experts to a group. The experts' opinions differ from popular opinion polls. Less competent experts can, however, adversely influence the final decision, even if the majority consists of high-quality experts. They argue that a larger number of specialists (experts) can be useful in complex environments where different specialisations are required to assess specific aspects of a decision.

When selecting experts, the most common methods use accreditation (if available), or a method based on experience and/or identification by collaborators. The method of experience, which considers the number of years worked or the number of specific experiences, is one of the oldest assuming that if someone performs in the area for a certain number of years, s/he is considered an expert (Shanteau et al., 2002).

The system of advisory bodies varies considerably from country to country and sector to sector. In most countries, there are permanent and ad-hoc advisory bodies. Permanent bodies make the system more stable but also less flexible in terms of its ability to respond to new problems. They often have a wide scope, long duration and expertise in a certain area; they monitor, estimate trends and collect data. These bodies are usually a part of the public administration. Our first group of experts consisted of permanent advisors/analysts. The selection process for experts was determined by the position of the staff at the ministries and included staff from the analytical units.

Ad-hoc bodies may lack a sufficiently broad basis for structural and broader analyses and are used to answer questions quickly. They consist of external specialists, who are part of a so-called partnership ensuring the participatory approach when designing EU-supported interventions. This partnership includes regional and local administration, the academy, non-governmental organisations and social partners. Their staffing is more diverse, inclusive and external (OECD, 2017). The second group, comprised of specialists, represents the ad-hoc advisory body. Our selection was based on two criteria: at least 10 years of experience and recommendation by several members of the partnership. This approach reflects the common selection procedure when external experts are involved in the decision-making process in Slovakia.

3 RESULTS

An important indication of judgement quality is consistency. Consistency in the AHP is measured via the Consistency Index (CI) and Consistency Ratio (CR). Research has confirmed that $CR \leq 0.1$ is acceptable for the AHP exercise (Chu and Kuang-Han, 2002; Franek and Kresta, 2014). Consistency was measured for

all respondents in both groups. The respondents with a high level of inconsistency were excluded from the sample. Over 100 experts were addressed. The overall number of completed replies was 45 (20 analysts and 25 specialists), which can be considered a very good response rate, but only 64% of those were consistent (14 analysts and 15 specialists) and could be further processed (Table 1). Therefore, the size of individual decision-making groups is not always optimal. This does not allow generalisation of our findings but provides sufficiently good indications to identify differences. It limits the risk of unwanted influence caused by inconsistent decision makers on the final decision.

Table 1 – Final Number of Replies in Both Groups

Topic	**Analysts**		**Specialists**	
	Consistent	Inconsistent	Consistent	Inconsistent
1. Employment and Labour Market	5	3	3	4
2. Education	5	2	4	3
3. Social Inclusion and Soc./Health Serv.	4	1	7	3
Total	14	6	15	10

Out of 29 consistent replies, six respondents provided replies for two topics (three from each group) and two assessed measures for all three topics (one from each group). However, consistent replies were provided by only one person for all three topics and by three respondents dealing with two topics (Table 2). This may suggest that a deeper knowledge of a certain area could be crucial for consistent decision-making. This assumption has not been found in the studied literature, but it offers the opportunity for future research.

Table 2 – Number of Replies and Consistent Replies Provided by the Groups

Decision makers	**1 Topic**		**2 Topics**		**3 Topics**	
	Replied	Consistent	Replied	Consistent	Replied	Consistent
Analysts	12	8	3	3	1	0
Specialists	16	11	3	0	1	1
Overall	28	19	6	3	2	1

Consistency of judgement (especially inconsistency on criteria and total inconsistency) was very similar in both the analyst and specialist groups (Table 3).

Table 3 – Consistency of Groups (AIJ) on Individual Criteria

Topic	Analysts	Specialists	Overall
Inconsistency on criteria	0.030	0.027	0.029
Inconsistency on relevance	0.081	0.032	0.056
Inconsistency on urgency	0.066	0.092	0.080
Inconsistency on feasibility	0.020	0.037	0.029
Total inconsistency	0.040	0.038	0.039

To process all collected inputs, aggregated individual judgement (AIJ) and aggregated individual priority (AIP) procedures can be applied. Both AIP and AIJ procedures conduct aggregation through the weighted arithmetic mean (WAM) or weighted geometric mean (WGM) methods, depending on the group size and structure (Bernasconi, Choirat and Seri, 2014). The AIJ procedure cannot guarantee the Pareto optimality axiom, which states that if all group members prefer alternative A1 to alternative A2, then the group should prefer A1 to A2 as well (Ossadnik, Schinke and Kaspar, 2016), so the AIP method was chosen for this research as more appropriate.

As the groups were consistent, the AIPs were calculated through both WAM and WGM. The WAM is provided in Table 4 to better demonstrate the priorities decided in the AHP. The total sum of the priority indexes in each topic equals one, so the preference is easily visible. The topics themselves were not compared as most of the respondents assessed the measures in only one of the topics. The prioritisation of measures resulted in a very similar set of priorities determined by both groups. The priorities set out for the employment and labour market were practically identical. There was a minor difference in priorities for education, and a slightly different order was assigned to priorities in social inclusion and social-health services but with minor differences (Table 4). When comparing priorities according to the individual criteria, the biggest differences appeared in the perception of feasibility of measures 2.2 and 2.3. The analysts saw the *Promotion of equal access* more feasible (0.29) than the specialists did (0.18) and, vice versa, the *Support of life-long learning* is considered more feasible by the specialists (0.39) than by the analysts (0.25). Measure 3.3, *Improving access to social security services and healthcare*, showed similar differences for two criteria. The relevance and feasibility were assessed high by the analysts (0.39 resp. 0.36) while the specialists assigned lower values to both criteria (0.22 resp. 0.29). The other values did not indicate substantial differences.

Table 4 – Weighted Arithmetic Mean of Aggregated Individual Priority / and Ranking for Individual Measures (by Group)

Topic – Measure	Analysts	Specialists	Overall
1. Employment and Labour Market			
1.1 Access to employment	0.29 / 1.	0.30 / 1.	0.29 / 1.
1.2 Modernisation of labour market institutions and services	0.28 / 2.	0.26 / 2.	0.28 / 2.
1.3 Promoting gender balanced labour market participation	0.20 / 4.	0.18 / 4.	0.19 / 4.
1.4 Support of adaptation	0.23 / 3.	0.26 / 2.	0.24 / 3.
2. Education			
2.1 Improving the quality of education	0.44 / 1.	0.37 / 2.	0.41 / 1.
2.2 Promoting equal access	0.35 / 2.	0.38 / 1.	0.36 / 2.
2.3 Support of life-long learning	0.21 / 3.	0.25 / 3.	0.23 / 3.
3. Social Inclusion and Social/Health Services			
3.1 Promoting active inclusion	0.18 / 4.	0.19 / 4.	0.19 / 4.
3.2 Support of socio-economic integration of MRC*	0.21 / 3.	0.27 / 2.	0.25 / 3.
3.3 Improving access to social security services and healthcare	0.37 / 1.	0.25 / 3.	0.29 / 1.
3.4 Promoting social integration of people at risk of poverty	0.24 / 2.	0.29 / 1.	0.27 / 2.

Notes: MRC* Marginalised Roma Communities

As the pilot exercise carried out in 2019 had the identical structure of measures for topics 2 and 3, it is possible to compare the priorities (Table 5). The original priorities did not highlight the quality of education to the same extent, and the main emphasis was on inclusion and life-long learning. We assume that the current change in the order of priorities may have been influenced by the global COVID-19 pandemic, and the long period of online education for Slovak schools focused the attention of the analysts on its quality (measure 2.1). A similar perception could be reflected in the higher priority of the measure related to the health sector (3.3). Both education and health are currently considered the main governmental priorities. The priorities set by the specialists highlight the same measures but more proportionally, with about the same weight. This provides approximately the same division of priorities in the social inclusion topic as in 2019 (Table 5).

Table 5 – Weighted Arithmetic Mean of Aggregated Individual Priority / and Ranking for Individual Measures in 2021 and 2019

Topic – Measure	Analysts 2021	Specialists 2021	Overall 2021	Overall 2019
2. Education				
2.1 Improving the quality of education	0.44 / 1.	0.37 / 2.	0.41 / 1.	0.26 / 3.
2.2 Promoting equal access	0.35 / 2.	0.38 / 1.	0.36 / 2.	0.46 / 1.
2.3 Support of life-long learning	0.21 / 3.	0.25 / 3.	0.23 / 3.	0.38 / 2.
3. Social Inclusion				
3.1 Promoting active inclusion	0.18 / 4.	0.19 / 4.	0.19 / 4.	0.17 / 4.
3.2 Support of socio-economic integration of MRC*	0.21 / 3.	0.27 / 2.	0.25 / 3.	0.30 / 1.
3.3 Improving access to social security services and healthcare	0.37 / 1.	0.25 / 3.	0.29 / 1.	0.28 / 2.
3.4 Promoting social integration of people at risk of poverty	0.24 / 2.	0.29 /1.	0.27 / 2.	0.24 / 3.

Notes: MRC* Marginalised Roma Communities

Based on the assumption of the economic theory that the rational decision maker is consistent, we measured consistency. The CRT is also confirmed as a substantial independent predictor of a group of rational thinking tasks (Toplak, West and Stanovich, 2014). The cognitive reflection score was thus measured together with consistency. The comparison of the consistency in the groups shows the same values, although the cognitive score is more than 20% higher for the analyst group (Table 6). Both groups scored high in the CRT, as the average score of common population was 4.2 (Sirota et al., 2021). It is known that people with a lower CRT are significantly more likely to be subject to overconfidence (Noori, 2016). In our experiment, the inconsistent respondents scored as high as the group of specialists (6.7) but their overconfidence was nearly twice as high as for the consistent respondents (2.6).

Table 6 – CRT Score and Overconfidence (by Group)

Main Group Features	Analysts	Specialists	Overall	Inconsistent respondents
CRT score	8.42	6.95	7.00	6.70
Overconfidence	1.58	1.90	1.38	2.60

The attempt to identify a correlation between consistency and both CRT score and the number of correct answers estimated by respondents was not successful (Table 7). Overconfidence was also calculated to test possible correlation with total consistency. The correlation could not be identified, either; first, because the overestimation for 80% of respondents was in the range (–2 to 2) – that is, the

sample was rather homogenous with only minor deviations. Second, the small sample, which complied with the size necessary for the AHP, was not sufficiently representative and large enough to detect a correlation. The results also suggest that the selection of experts and further selection of consistent experts, keeps the CRT score high and overconfidence at a similar level.

Table 7 – Correlation Matrix for Total Consistency, Confidence, and CRT Score (by Group)

	Analysts	**Specialists**	**Overall**
Total consistency (TC)	0.040	0.038	0.039
Correlation			
TC - CRT score	-0.285	-0.511	-0.428
TC - estimated correct CRT answers	0.202	-0.361	-0.139
TC - overconfidence	0.557	0.177	0.335

4 CONCLUSION

The authors used this unique opportunity to combine practical policy exercises with scientific research on high-stakes decisions through the use of AHP. High-stakes decision makers are not easily accessible for research, so their participation due to the support and involvement of the MIRDI is highly appreciated. The newly tested approach for strategic decision-making with the support of internal and external experts was beneficial. The original 'foggy' decision-making process in high-stakes strategic public policies was replaced by a transparent and operational mechanism utilising up-to-date research methodologies. This proved to be a functional managerial innovation that made it possible to carry out a participatory decision-making process with relevant stakeholders.

In addition, for the first time ever, the decision-making of expert groups was compared, and it was confirmed that groups of analysts and specialists are similar as regards their basic measured characteristics. The groups were equally (in)consistent, with slightly higher CRT score for analysts and slightly lower overconfidence compared to specialists. However, comparison of their preferences showed some differences and confirmed slightly different set of the priorities. We can assume that higher priorities assigned to quality of education and health measures by analysts were quite likely driven by the focus of governmental priorities highlighted by the COVID-19 pandemic. Comparison of the priorities set out in 2019 and 2021 reflects the changes in 2 "COVID" years.

These results indicate that the selection process for experts is very important. The consistent respondents in both groups produced similar results. The assumed relationship between consistency and CRT score and/or overconfidence was not

detected, as the consistent decision makers proved to be well calibrated. This suggests the possible application of consistency as a calibration tool, but this requires further research. Similar conclusions can be drawn from the attempt of a few respondents to assess more than one topic, which suggests that high expertise in a topic, leads to higher consistency, but if more topics are selected the inconsistent replies are more frequent. This hypothesis needs verification in a well-designed experiment. Although for our experiment the size and composition of the expert groups was in line with the general practice in Slovakia, it did not enable generalisation of our findings. The main limitations of our research are the small sample of decision makers; while this sample size complies with the requirements of AHP, it is not sufficient to confirm the statistical validity of our findings. As it is not practically feasible to increase the size of the groups, future research could benefit from the further repetition/application of the group decision-making, which is specifically meant for high-stakes decision-making in public policy.

REFERENCES

Bačová, V., 2010. *Rozhodovanie a usudzovanie. Pohľady psychológie a ekonómie I.* Bratislava: Ústav experimentálnej psychológie SAV.

Baláž, V., 2009. *Riziko a neistota. Úvod do behaviorálnej ekonómie a financií.* Bratislava: VEDA.

Baláž, V., Dokupilová, D. and Filčák, R., 2021. Participatory multi-criteria methods for adaptation to climate change. *Mitigation and Adaptation Strategies for Global Change*, 26(4), pp.1-22. DOI: 10.1007/s11027-021-09955-4.

Bazerman, M.H. and Moore, D.A., 2013. *Judgment in Managerial Decision Making*. Wiley.

Bernasconi, M., Choirat, C. and Seri, R., 2014. Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence. *European Journal of Operational Research*, 232(3), pp.584-592. DOI: 10.1016/j.ejor.2013.06.014.

Berrittella, M., Certa, A., Enea, M. and Zito, P., 2011. An Analytic Hierarchy Process for the Evaluation of Transport Policies to Reduce Climate Change Impacts. [FEEM Working Paper No. 12.2007] *SSRN Electronic Journal*. Available at SSRN: <https://ssrn.com/abstract=962379> DOI: 10.2139/ssrn.962379.

Bhushan, N. and Rai, K., 2007. *Strategic Decision Making: Applying the Analytic Hierarchy Process.* Springer Science & Business Media.

Burgman, M.A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L. and Twardy, Ch., 2011. Expert Status and Performance. *PLoS ONE,* 6(7), e22998. DOI: 10.1371/journal.pone.0022998.

Čavojová, V. and Hanák, R., 2014. Assessing fluid rationality and its relation to cognitive styles. In: T. Ahram, W. Karwowski and Marek, T., eds. *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics*. Krakow: AHFE. pp.142-153.

Chu, P. and Kuand-Han, L. J., 2002. Note on consistency ratio. *Mathematical and Computer Modelling*, 35, pp.1077-1080.

Danesh, D., Ryan, M. and Abbasi, A., 2017. A Systematic Comparison of Multi Criteria Decision Making Methods for the Improvement of Project Portfolio Management in Complex Organisations. *International Journal of Management and Decision Making*, 16(1), pp.280-320.

Dawes, R., 1979. The Robust Beauty of Improper Linear Models in Decision Making. *American Psychologist*, 34, pp.571-82.

Dokupilová, D., Baláž, V., Kurincová Čavojová, V., Ballová Mikušková, E. and Gombitová, D., 2020. Identifying major policy challenges and policy interventions via expert methods application of the Delphi and AHP methods in preparation of the Partnership Agreement for the Slovak Republic in period 2021-2027. *Národohospodářský obzor*, 20(3), pp.361-377. DOI: 10.2478/revecp-2020-0017.

Dunning, D., Johnson, K., Ehrlinger, J. and Kruger, J., 2003. Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), pp.83-87. DOI: 10.1111/1467-8721.01235.

Emrouznejad, A. and Marra, M., 2017. The state-of-the-art development of AHP (1979–2017): A literature review with a social network analysis. *International Journal of Production Research*, 55(22), pp.6653-6675.

Franek, J. and Kresta, A., 2014. Judgment scales and consistency measure in AHP. *Precedia Economics and Finance*, 12, pp.164-173.

Frederick, S., 2005. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), pp.25-42. DOI: 10.1257/089533005775196732.

Gilmour, P. and Corner, J.L., 1998. The Role of the Expert's Decision Making Skills in Management. *33rd Annual Operational Research Society of New Zealand,* pp.195-204.

Hutton, R.J.B. and Klein, G., 1990. Expert decision making. *Systems Engineering*, 2(1), pp.32-45.

Ishizaka A. and Labib A., 2011. Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications*, 38(11), pp.14336-14345.

Juanchich, M., Sirota, M. and Bonnefon, J. F., 2020. Anxiety-induced miscalculations, more than differential inhibition of intuition, explain the gender gap in cognitive reflection. *Journal of Behavioural Decision Making*, 33, pp.427-443. DOI: 10.1002/bdm.2165.

Kahneman, D., 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kruger, J. and Dunning, D., 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence led to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), pp.1121-1134.

Larrick, R.P. and Feiler, D.C., 2015. Expertise in decision making. In: G. Keren, G. Wu, eds. 2015. *The Wiley Blackwell Handbook of Judgment and Decision Making*. John Wiley & Sons. pp.696-721.

Lockton, D., 2012, Cognitive biases, heuristics and decision-making in design for behaviour change. [Working Paper] *SSRN Electronic Journal*. Available at SSRN: <https://ssrn.com/abstract=2124557>. DOI: 10.2139/ssrn.2124557.

Mu, E. and Pereyra-Rojas, M., 2017. AHP Models with Sub-criteria. In: E. Mu and M. Pereyra-Rojas, eds., 2017. *Practical Decision Making*. Cham: Springer. pp.45-56.

Noori, M., 2016. Cognitive reflection as a predictor of susceptibility to behavioural anomalies. *Judgment and Decision Making*, 11(1), pp.114-120.

OECD, 2017. *Policy Advisory Systems: Supporting Good Governance and Sound Public Decision Making*. Paris: OECD Public Governance Reviews, OECD Publishing. DOI: 10.1787/9789264283664-en.

Oppenheimer, D.M. and Thomson, K.S., 2016. Investigation of an alternate form of the Cognitive Reflection Test. *Judgment and Decision making*, 11(1), p.99.

Ossadnik, W., Schinke, S. and Kaspar, R.H., 2016. Group Aggregation Techniques for Analytic Hierarchy Process and Analytic Network Process: A Comparative Analysis. *Group Decision and Negotiation*, 25(2), pp.421-457. DOI: 10.1007/s10726-015-9448-4.

Pennycook, G., Ross, R. M., Koehler, D. J. and Fugelsang, J. A., 2017. Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence. *Psychonomic Bulletin and Review*, 24(6), pp.1774-1784.

Peregrin, S. and Karahuta, M., 2014. Skupinový analytický hierarchický proces a jeho využitie v manažmente spoločnosti/Group analytic hierarchy process and its usage in management of company. *Mladá Veda*, 2(3), p.60.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M.A. and Hamilton, J., 2016. The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, 29(5), pp.453-469.

Saaty, T., 2008. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), pp.83-98.

Saaty, T.L. and Özdemir, M.S., 2014. How Many Judges Should There Be in a Group? *Annals of Data Science*, 1(3-4), pp.359-368.

Saaty, T.L. and Peniwati, K., 2013. *Group decision making: drawing out and reconciling differences*. RWS publications.

Shanteau, J. and Stewart, R., 1992. Why Study Expert Decision Making? Some Historical Perspectives and Comments. *Organisational Behaviour and Human Decision Processes*, 53, pp.95-106.

Shanteau, J., Weiss, D., Thomas, R.P. and Pounds, J.C., 2002. Performance based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, 136(2), pp.253-263.

Sičáková-Beblavá, E., 2015. *Behaviorálne základy verejnej politiky*. Bratislava: Univerzita Komenského v Bratislave.

Sinayev, A. and Peters, E., 2015. Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology*, 6, p.532. DOI: 10.3389/fpsyg.2015.00532.

Sirota, M. and Juanchich, M., 2018. Effect of response format on cognitive reflection: Validating a two- and four-option multiple choice question version of the Cognitive Reflection Test. *Behaviour Research Methods,* 50, pp.2511-2522.

Sirota, M., Dewberry, Ch., Janchich, M., Valuš, L. and Marshall, A.C., 2021. Measuring cognitive reflection without maths: Development and validation of the verbal cognitive reflection test. *Journal of Behavioural Decision Making*, 34, pp.322-343. DOI: 10.1002/bdm.2213.

Sleboda, P. and Sokolowska, J., 2017. Measurements of Rationality: Individual Differences in Information Processing, the Transitivity of Preferences and Decision Strategies. *Frontiers in Psychology*, 8, p.1844. DOI: 10.3389/fpsyg.2017.0184

Šoltés, V. and Gavurová, B., 2014. The functionality comparison of the health care systems by the analytical hierarchy process method. *Ekonomie a Management,* 17(3), pp.1-18.

Subramanian, N. and Ramanathan, R., 2012. A review of applications of Analytic Hierarchy Process in operations management. International Journal of Production Economics, 138, pp.215-241. DOI: 10.1016/j.ijpe.2012.03.036

Tetlock, P.E., 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press.

Toplak, M.E., West, R.F. and Stanovich, K.E., 2011. The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39(7), pp.1275-1289. DOI: 10.3758/s13421-011-0104-1.

Toplak, M.E., West, R.F. and Stanovich, K.E., 2014. Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20(2), pp.147-168. DOI: 10.1080/13546783.2013.844729.

Tversky, A., 1969. Intransitivity of preferences, *Psychological Review*, 76, pp.31-48.

Vargas, R.V., 2010. Using the analytic hierarchy process (AHP) to select and prioritize projects in a portfolio. In: Project Management Institute (PMI®), *Global Congress 2010-North America*. 12 October 2010. Washington, DC. Newtown Square, PA: Project Management Institute. Available at: <https://www.pmi.org/learning/library/analytic-hierarchy-process-prioritize-projects-6608> [10 December 2021].

Velasquez, M. and Hester, P.T., 2013. An Analysis of Multi-Criteria Decision-Making Methods. *International Journal of Operations Research*, 10(2), pp.55-66.

Vidal, L.M., Marle, F. and Bocquet, J.C., 2011. Using a Delphi process and the Analytic Hierarchy Process (AHP) to evaluate the complexity of projects. *Expert Systems with Applications*, 38(5), pp.5388-5405.

Wagenaar, W.A. and Keren, G.B., 1985. Calibration of probability assessments by professional blackjack dealers, statistical experts, and lay people. *Organizational Behaviour and Human Decision Processes*, 36(3), pp.406-416.

Wallenius, J., Dyer, J.S., Fishburn, P.C., Steuer, R.E., Zionts, S. and Deb, K., 2008. Multiple criteria decision-making, multi-attribute utility theory: Recent accomplishments and what lies ahead. *Management Science*, 54(7), pp.1336-1349.

Weinstein, B.D., 1993. What is an expert? *Theoretical Medicine and Bioethics*, 14(1), pp.57-73.

Welsh, M.B., Burns, N.R. and Delfabbro, P.H., 2013. The Cognitive Reflection Test: How much more than numerical ability? In: M. Knauff, M. Pauen, N. Sebanz and I. Wachsmuth, ed., *35th Annual Conference of the Cognitive Science Society*. Berlin, Germany, July 31-August 3, 2013. Austin, TX: Cognitive Science Society. pp.1587-1592.

Wu, Z. and Tu, J., 2020. Managing transitivity and consistency of preferences in AHP group decision making based on minimum modifications. *Information Fusion*, 67, pp.125-135. DOI: 10.1016/j.inffus.2020.10.012.

Zyoud, S.H. and Fuchs-Hanusch, D., 2017. A bibliometric-based survey on AHP and TOPSIS techniques. *Expert Systems with Applications*, 78(C), pp.158-181. DOI: 10.1016/j.eswa.2017.02.016.

ABOUT AUTHORS

Dagmar Gombitová[0000-0001-6204-9238] (D.G.) – Ing. Mgr., PhD. Student at the Institute for Forecasting, Centre of Social and Psychological Studies, Slovak Academy of Sciences (CSPS SAS), Bratislava, Slovakia, and Public Policy Institute, Faculty of Social and Economic Sciences, Comenius University Bratislava, Slovakia, e-mail: dagmar.gombitova@savba.sk.

Dušana Dokupilová[0000-0001-9813-2201] (D.D.) – RNDr. PhD., Researcher at the Institute for Forecasting, Centre of Social and Psychological Studies, Slovak Academy of Sciences (CSPS SAS), Bratislava, Slovakia, e-mail: dusana.dokupilova@savba.sk.

AUTHOR CONTRIBUTIONS

Conceptualization: D.G.; Methodology: D.G. and D.D.; Software, D.D.; Validation, D.D.; Formal analysis, D.D.; Investigation, D.G.; Resources, D.G.; Data curation: D.D.; Original draft preparation: D.G.; Review and editing: D.D.

CONFLICTS OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.