Досліджено відомі шкали (критерії) для оцінювання компетентності експертів у сфері технічного регулювання із застосуванням метода аналітичної ієрархії (МАІ) і моделі Раша. Представлено основні особливості побудови математичної моделі Раша. Розглянуто результати аналізу шкал для оцінювання компетентності експертів у сфері технічного регулювання на конкретному прикладі анкетованих фахівців з вимірювання часу і частоти. Результати опрацьовано за допомогою спеціалізованих програмних засобів "Компетентність МАІ 1.1" (Україна) і MINISTEP 4.0.1 (США).

Проведено порівняльний аналіз результатів з метою визначення ефективності шкал оцінювання. Отримані результати показали можливість застосування моделі Раша для аналізу шкали оцінювання компетентності експертів у сфері технічного регулювання. Аналіз отриманих за багатомірною моделлю Раша результатів показав, що обрана шкала критеріїв для експертів відповідає встановленим до моделі Раша вимогам. Отримані дані вимірювань за цією моделлю дозволяють розрахувати встановлені статистики як для критеріїв, так і для експертів, що оцінені.

Порівняльний аналіз результатів, отриманих за допомогою МАІ та за моделлю Раша, показав збіжність, придатність та наявність кореляції отриманих значень за експертами. Лише два з двадцяти одного (9,5 %) оцінених експертів має дані, які непридатні для аналізу за моделлю Раша, що свідчить про низьку компетентність. МАІ у меншій ступені дозволяє враховувати думки менш компетентних експертів, ніж із застосуванням моделі Раша. Про це свідчить менший коефіцієнт компетентності для МАІ, ніж при застосуванні моделі Раша. МАІ та модель Раша доцільно застосувати в різних сферах діяльності як корисний інструмент для порівняльної оцінки компетентності технічних експертів на основі об'єктивних даних за встановленими критеріями

Ключові слова: аналітична ієрархія, модель Раша, характеристична крива, компетентність експертів, шкала оцінювання, програмний засіб

A COMPARATIVE ANALYSIS OF THE ASSESSMENT RESULTS OF THE COMPETENCE OF TECHNICAL EXPERTS BY METHODS OF ANALYTIC HIERARCHY PROCESS AND WITH USING THE RASCH MODEL

O. Velychko
Doctor of Technical Sciences, Professor, Director
Scientific and Production Institute of
Electromagnetic Measurements
State Enterprise "All-Ukrainian State Scientific
and Production Centre for Standardization,
Metrology, Certification and Protection of
Consumer", (SE "Ukrmetrteststandard")
Metrolohychna str., 4, Kyiv, Ukraine, 03143
E-mail: velychko@hotmail.com

T. Gordiyenko
Doctor of Technical Sciences, Associate
Professor, Head of Department
Department of standardization,
conformity assessment and quality
Odessa State Academy of
Technical Regulation and Quality
Kovalska str., 15, Odessa, Ukraine, 65020
E-mail: t_gord@hotmail.com

1. Introduction

For the acceptance of reasonable decisions in any spheres of activity, it is necessary to lean against experience, knowledge and intuition of specialists. To that end, group expert evaluations are conducted, carefully approaching the selection of specialists in a certain field of activity. The importance of attracting highly skilled professionals to identify and resolve problem issues in any field of activity is beyond doubt.

Particularly relevant is the use of methods of group expert evaluation for the implementation of reforms in the field of technical regulation, in particular, in metrology. In this area, specialists work in different directions, in particular standardization, metrology, conformity assessment, etc. In order to obtain reliable estimates of group expert evaluation in any field of activity, it is needed to select the most optimal method. To this end, a comparative analysis of the assessment of the suitability of the methods is used. The choice of the most optimal method for group expert evaluation and a comparative analysis of suitability for application to increase the efficiency of complex systems were the subject of previous studies by the authors [1–5].

2. Literature review and problem statement

Expert evaluation techniques are based on the use of knowledge of skilled experts – experts in the subject field. To assess the experts, there are a number of characteristics that are used to select both specific experts and expert groups: the coefficient of competence (CC), the coefficient of concordance, the coefficient of reliability of expert assessments, etc. A detailed analysis of the most commonly used expert methods, advantages and disadvantages of each has been presented in [1–5].

For research of complex objects or systems, the method of analytical hierarchy (Analytic Hierarchy Process, AHP) has been selected [6–8]. This method is a mathematical instrument of the systems approach to complex problems of decision-making, allows in the interactive mode to find such variant (alternative) that best comports both with its understanding of the problem and with requirements in relation to its solution [3, 4]. Although, the base method of AHP [6–8] and its modifications [9–12] are applicable only in case of a small number of the set of alternatives and do not give an opportunity to combine different opinions of groups of experts.

In recent decades, Rasch mathematical model [13] is widely used to create new or to view existing scales. Narrow definition of the Rasch model is a method of transforming the received primary data into an interval scale of natural logarithms. At the same time, the primary data in the model are not considered in the process of logarithmic transformation. This model ensures that valid results are obtained through the use of statistics of adequacy, diagnostic information, presents test parameters on a single common linear scale, which helps in the criterion-oriented interpretation of the data.

Scientific publications for the Rasch model cover quite a lot of spheres of activity [14–18]. However, there are practically no publications covering the field of technical regulation, in particular metrology.

In [19], the Rasch model has been defined as a comparison of the results of natural logarithms studied on the scale. The mathematical side and the very theory of G. Rasch have been successfully developed in [14]. If the data correspond to the Rasch model, then as a result, they are presented on an interval scale that is resistant to the loss of some primary data. Therefore, the model is a method of objective scaling of data. Several software tools, including the most commonly used software MINISTEP 4.0.1 (USA) [20], have been developed to allow for the necessary calculations based on the Rasch model, as well as to provide an appropriate assessment of the suitability of the data for the model used.

3. The aim and objectives of the study

The purpose of the work is to identify the most effective method and means for expert evaluation suitable for assessing the competence of technical experts in the field of technical regulation.

To achieve the aim, the following objectives were accomplished:
– to carry out the analysis of known scales (criteria) for assessing the competence of experts in the field of technical regulation with the use of AHP and the Rasch model;
– to evaluate the possibility of applying the Rasch model for analyzing the scale of expert assessment in the field of technical regulation;

– to conduct a comparative analysis of the results obtained by the use of the AHP and the Rasch model, and to determine the effectiveness of the scales with the indication of the obtained characteristics of the Rasch model.

4. Materials and methods of researching the methods of assessment of competence of technical experts

4. 1. Method of assessment of competence of technical experts based on the Analytic Hierarchy Process

The AHP [6], used in the case of a small number of given alternatives, when the decision-maker (DM) efforts aim at comparing only the given alternatives is widely known. The task for the AHP method is that with a known general aim (or sub-aim) for a solution to a problem, N criteria for evaluating alternatives and n alternatives, it is needed to choose the best alternative.

The AHP is designed to determine the optimal variant (from several), taking into account many criteria of different nature and allows you to use a variety of criteria (quantitative, qualitative, numerical with different dimensionality, etc.) when comparing. Its maximum effectiveness is manifested in the search for solutions to complex problems requiring a systematic approach and involving a large number of experts [1].

Basic phases for implementation of the AHP method are:
– to structure a task as a hierarchical structure with a few levels (aims-criteria-alternatives);
– to execute pairwise comparisons of items of every level and transform the results of comparisons in numbers by means of the special scale of relative importance;
– to calculate the coefficients of weight for the items of every level and check the consistency of judgments of DM;
– to carry out the calculation of the quantitative indicator of quality of each of alternatives and determine the best alternative.

The resulting global priorities for each expert are ranked in order of increasing global priority values (G_{ni}). An expert who has received the maximum value of a global priority is considered the most competent [3].

The mathematical apparatus of the AHP is described in detail in [4], which also presents the results of evaluating the competence of experts in metrology with this method.

4. 2. Method of assessment of competence of technical experts with using the Rasch model

Analysis of the data by the Rasch model gives a number of details to verify that the added results in the data are justified. This is called the matching test between the data received and the selected model. If the data adequately correspond to the goals of the chosen model, the analysis also linearizes the overall assessment, which is limited to the level 0 and the maximum estimate for certain objects under investigation. The measure of Rasch is a linear value on the additive scale representing a hidden variable.

The researcher can use the analysis of the Rasch model to check the degree to which the estimation and summation for this model are reliable in the data obtained. Within the framework of the Rasch model, the relationship between the probability of success to the item and the hidden feature is described by a special function. This function is called item characteristic curve (ICC) or item response function (IRF), which has an S-shape (Fig. 1). The function shows the link

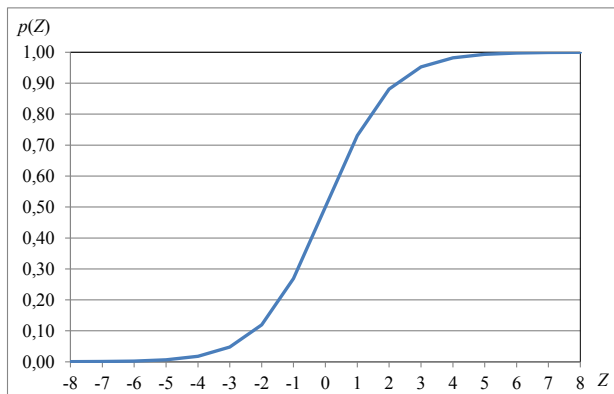between the overall assessment of the test and the assessment of the location of the subject (expert).

Fig. 1. Item characteristic curve

In the Rasch model, the index of distribution of the subject, including the measurement error, is used instead of reliability indicators. The magnitude of the measurement error is not uniform over the test range, but is usually greater for more extreme points (low and high). The scale of the successive response of each subject to each item (the Rasch scale) has interval-scale properties. Interval scales are known for even intervals between two gradations: certain numeric values indicate how much more the attribute of an item is present.

The Rasch model suggests that the probability of approving any category of responses to an object depends entirely on the subject's ability and complexity of the object. That is, no other attribute of subjects or objects determines the possibility of measurement in order to determine the likelihood of approval of the response [21].

Rasch linear scales are initially expressed in units within 1, but can be redistributed in accordance with normal scaling, from 0 to 100, while maintaining aggregate additivity. The Rasch model also estimates the calculation error in each level as standard measurement errors. The error is always greater on the upper and lower ends of the scale as the Rasch model is not limited to the boundaries, but measures from the middle of the range of values and provides infinity in both directions. Measurement is better when the average values of the items lie closer to the average values of the scale, that is, the real assessment is more uncertain when approaching the bounds of the scale [22].

The final stages of obtaining the characteristics of the Rasch model based on the best solutions are:
– uniform arrangement of values of items (equality of scale steps);
– reduced measurement error (high accuracy);
– the probability and unlikelihood (suitability) of items and qualities of the subject expected from the model;
– overall reliability (noise – excessive unpredictability of data, possibly due to excessive randomness or multidimensionality);
– simplicity;
– conformity of the nature of the measured items.

Special characteristic – logit is a key element for the probabilistic Rasch model [13]. Logit is the probability log unit – the unit of measurement used in the Rasch model for calibrating the items and measuring the subjects by the

hidden variable. That is, the logarithmic transformation of the ratio of the probabilities of the correct and incorrect response or the probabilities of neighboring items on a certain scale. Measurement means the location (usually in logits) by the hidden variable.

The logit of number p – probability, is determined by the formula:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = -\log\left(\frac{1}{p}-1\right). \tag{1}$$

The value of $p/(1-p)$ is the corresponding coefficient and the probability logit is the logarithm of odds.

Infit and Outfit statistics are the most widely used determinative statistics of the Rasch model. The Infit statistics are more critical when the scale of the item is close to the subject's scale, and the Outfit statistics are more critical if the indicators at the extreme limit of the scale are not metrics of the subject. Rasch charts and tables use normalized unweighted averages, so that the graphs are symmetric centered to zero [20].

Infit statistics are statistically weighted data or more sensitive compliance statistics that focus on the overall performance of an item or subject, that is, the weighted average of the standardized standard deviation of observation from the expected one (normalized mean square). Outfit statistics are sensitive to statistics that cover rare or unexpected events. This is the average of the squared values of standardized performance deviations from the expected performance.

5. Assessment of the competence of technical experts by methods of Analytic Hierarchy Process and with using the Rasch model

The papers [3–5] describe methods for evaluating the competence of experts (ECE) using different methods. For all these methods, the same criteria of the ECE that set a certain scale are used.

For the implementation of the indicated and described methods of AHP and with using the Rasch model, the following criteria of the ECE in the field of technical regulation, in particular, metrology, were applied [1, 3, 23]:

K_1 – education and scientific level in the field of technical regulation;

K_2 – overall experience;

K_3 – experience in the field of technical regulation;

K_4 – experience as an expert in the field of technical regulation;

K_5 – work status.

For the implementation of the AHP method, the values of matrices of pairwise comparisons (MPC) of criteria with normalized priority vectors for the selected criteria of the ECE and the weight coefficients for the selected criteria of the ECE [3] were determined. For the proposed criteria of the ECE, the largest number of MPC of criteria was actually $\lambda_{\max}=5.35$. Checking the consistency of the output data by the obtained consistency index $I_c=0.09$ and the consistency ratio $C_d=0.07$ showed that the ratio meets the requirements of consistency ($C_d \leq 0.1$). This demonstrates the consistency of the established criteria of the ECE [1, 3].

6. Example of evaluation of expert's competence for measurement of time and frequency

To assess the competence of experts in metrology, a survey was conducted using a specially developed questionnaire based on the criteria of the ECE. The competence of 21 experts on time and frequency measurements was evaluated. Of 21 experts involved, 16 (76 %) represented state-owned enterprises of the technical regulation system, 5 (24 %) were other enterprises.

The data on these experts obtained by questionnaire were processed using specialized software "Competence AHP 1.1" (Ukraine) [3] and MINISTEP 4.0.1 (USA) [20], based on the above described methods.

The results of the assessment of the competence of technical experts using the software "Competence AHP 1.1" (Ukraine), implemented by the AHP, are shown in Fig. 2.

The received primary data on these experts were processed using the software MINISTEP 4.0.1 [20], which implements the Rasch model. The results of the transformation of the input primary data by the items (criteria) and by the subjects (experts) in the Rasch measurement are shown in Table 1, 2 respectively.

The results of the measurement by items and by subjects are presented in logits in descending order and are shown in Table 1 and 2. Measurement error is based on the Rasch model, that is, the estimated value, which, when added and subtracted from the measurement in logits, gives a minimum distance before the difference becomes significant.
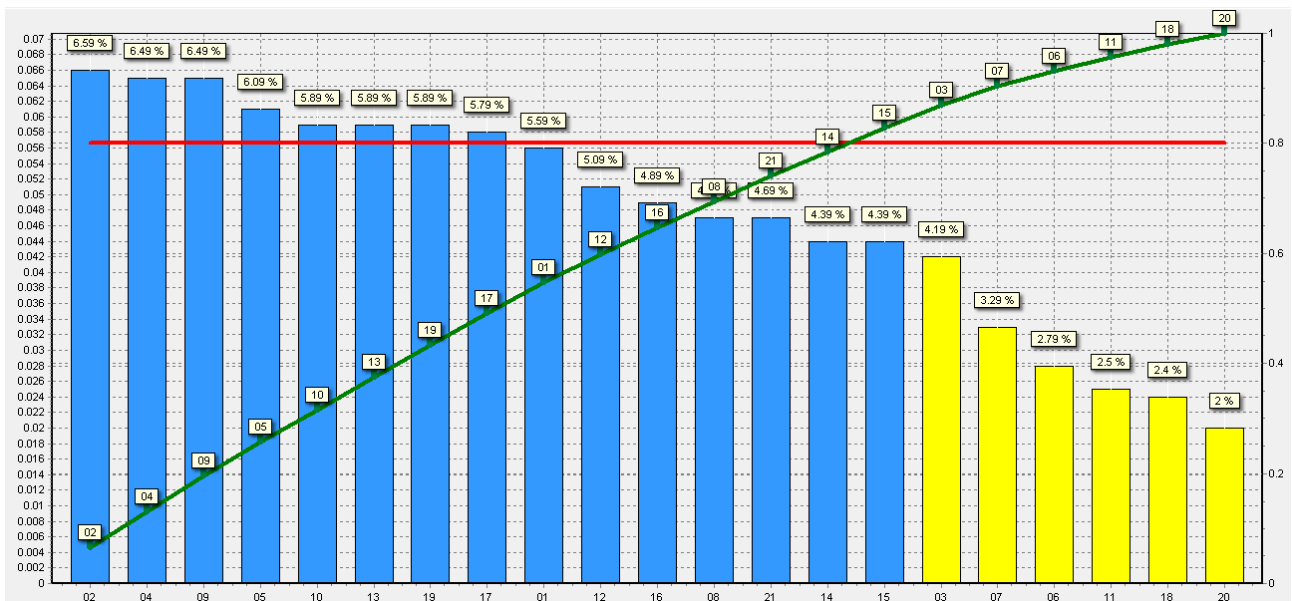

Fig. 2. Results of evaluation of expert's competence in metrology by AHP

Table 1

Results of conversion of data according to the ECE criteria

ECE criterion	Total score	Measurement	Measurement error	Infit statistics		Outfit statistics		Correlation coefficient	
				MNSQ	ZSTD	MNSQ	ZSTD	received	expected
K5	85	1.31	0.15	1.76	2.10	1.72	1.90	0.42	0.54
K6	174	−1.20	0.23	1.55	1.30	1.44	1.00	0.36	0.41
K2	145	−0.,22	0.16	0.99	0.10	0.98	0.00	0.76	0.52
K3	161	−0.68	0.18	0.46	−2.20	0.47	−1.80	0.74	0.50
K1	105	0.80	0.16	0.33	−2.50	0.28	−2.80	0.00	0.50
Average value	134.00	0.00	0.18	1.03	0.00	0.94	−0.10	–	–
Standard deviation	33.70	0.93	0.03	0.52	1.20	0.48	1.10	–	–

Table 2

Results of data conversion concerning experts

Expert	Total score, Z_i	Measu- rement	Measu- rement error	Infit statistics		Outfit statistics		Correlation coefficient	
				MNSQ	ZSTD	MNSQ	ZSTD	received	expected
18	28	0.25	0.31	2.43	2.10	2.77	2.30	0.20	0.78
21	36	1.25	0.41	1.46	0.80	2.34	1.40	0.44	0.78
10	32	0.68	0.35	1.53	0.90	1.29	0.60	0.94	0.78
20	25	−0.04	0.31	1.45	0.90	1.44	0.90	0.33	0.79
08	31	0.57	0.33	1.38	0.80	1.18	0.50	0.58	0.78
17	35	1,09	0.39	1.24	0.60	1.02	0.30	0.68	0.79
12	31	0,57	0.33	1.20	0.50	1.00	0.20	0.94	0.78
11	23	−0,24	0.32	1.06	0.30	1.10	0.40	0.81	0.80
14	34	0.94	0.38	0.56	−0.50	1.06	0.30	0.78	0.79
07	33	0.81	0.36	0.98	0.20	0.94	0.20	0.66	0.79
04	35	1.09	0.39	0.97	0.20	0.84	0.10	0.94	0.79
09	35	1.09	0.39	0.97	0.20	0.84	0.10	0.94	0.79
16	32	0.68	0.35	0.96	0.20	0.82	0.00	0.70	0.78
01	34	0.94	0.38	0.78	−0.10	0.64	−0.20	0.93	0.79
03	28	0.25	0.31	0.56	−0.80	0.64	−0.50	0.86	0.78
15	28	0.25	0.31	0.59	−0.70	0.55	−0.80	0.94	0.78
05	34	0.94	0.38	0.57	−0.40	0.55	−0.40	0.97	0.79
06	30	0.46	0.33	0.52	−0.80	0.48	−0.80	0.86	0.78
13	33	0.81	0.36	0.42	−0.80	0.43	−0.70	0.98	0.79
02	37	1.42	0.42	0.40	−0.70	0.33	−0.70	0.91	0.77
19	36	1.25	0.41	0.32	−0.90	0.30	−0.80	0.91	0.78
Average value	31.90	0.72	0.36	0.97	0.10	0.98	0.10	–	–
Standard deviation	3.70	0.43	0.03	0.49	0.80	0.60	0.80	–	–

The columns of Infit and Outfit statistics contain parameters that characterize the matching of the data of the Rasch model:

– MNSQ – the value that characterizes the level of randomness of the results or the discrepancy between the measurement model data;

– ZSTD – standardized MNSQ values, that is, the probability of mean-square-statistics, expressed as z-statistics (mean-square deviation).

MNSQ is also referred to as a relative xi-square or normalized xi-square.

Weighted average statistics of conformity are the xi-square statistics divided by its degrees of freedom. For the probability $p \leq 0.05$ (two-way distribution), ZSTD>|1.96|. The most expected values for MNSQ are near 1.0. The most qualitative and relevant values are MNSQ values ranging from 0.5 to 1.5. Values below 1.0 indicate that the data are either too predictable, or excessively predictable,

or overestimated model data. Values above 1.0 indicate too unpredictable data or underestimated model data. Values greater than 1.5 indicate uncertainty and "noise" (excessive unpredictability of data) in the input data, values less than 0.5 are also undesirable because they indicate an "information overload" of an item. The MNSQ values from −2.0 to +2.0 are acceptable. The values of MNSQ for a module larger than 2.0 are considered to be non-conforming to the measurement model and cannot be used in the analysis of the results. The analysis begins with questions of high MNSQ value.

The obtained MNSQ values for the criteria for Infit statistics range from 0.32 to 2.43, and for Outfit statistics from 0.55 to 1.86. This indicates that all these values are acceptable for the analysis by the Rasch model. Only for the criterion ECE K5, the values of the Infit and Outfit statistics are respectively 1.99 and 1.86, which indicates the presence of "noise" in the input data. In view of this, it is considered

expedient to clarify the points of the expert assessment scale only by this ECE criterion.

The obtained MNSQ values for experts for Infit statistics range from 0.22 to 1.53, except for 2.43 for expert 18, and for Outfit statistics from 0.30 to 1.44, except for 2.34 for expert 21 and 2.77 for expert 18. This indicates that all of these values are acceptable for the analysis of the Rasch model, in addition to the data for experts 18 and 21. Given this, it is expedient to remove the data on these experts from further analysis.

The correlation coefficient (may take values from –1 to +1) is considered as a measure of reliability and validity, and is used to identify, refine, and possibly exclude poorly matched items. The standard deviation is the mean square root of the difference between a sample of values and a mean. The obtained correlation coefficient for the criterion ECE K1 is equal to 0, and for the expert 18 is 0.20, which indicates a very small correlation of the corresponding data.

Using the software MINISTEP 4.0.1 (USA), graphical reports were also obtained: characteristic curves, information functions, etc. In Fig. 3, characteristic curves of all evaluated experts for all items, the analysis of the mutual placement of which helps to improve the evaluation as a system of criteria were constructed. In this case, most curves are concentrated in the middle and lower than average complexity. Characteristic curves practically uniformly fill the entire interval from –4.7 to +4.7 logits with the maximum allowable range from –5 to +5 logits. This indicates a fairly good agreement between the criteria of the ECE established for the evaluation of experts.
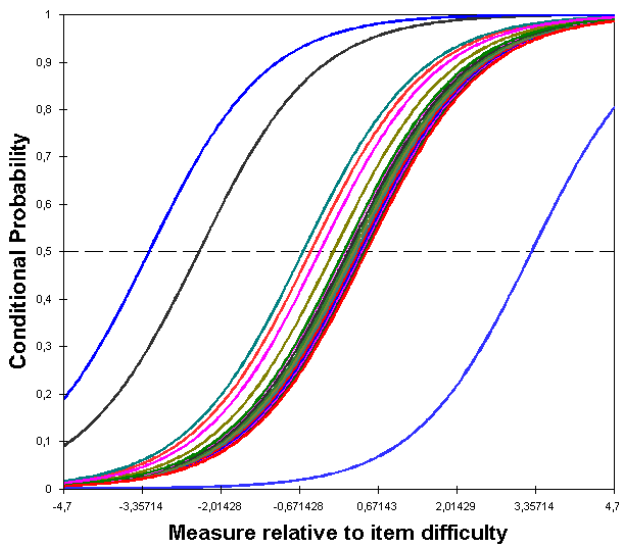

Fig. 3. Characteristic curves for evaluated experts, built using software MINISTEP 4.0.1

For each ECE criterion and evaluation as a whole, you can get a graphical representation of the correspondence of the data of the selected model (Fig. 4).

The obtained data indicate the presence of a correlation with the data for the selected model. Converted data for the evaluated experts according to the established criteria are shown in Fig. 4. This allows us to clearly show the ranking of experts based on the results of the application of the Rasch model for all established ECE criteria.
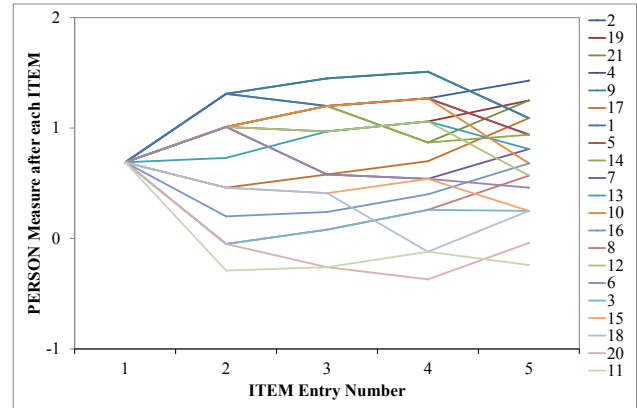

Fig. 4. Converted data on evaluated experts according to the criteria set by the ECE

7. Discussion of the results of the assessment of the competence of technical experts

In order to compare the results obtained with the use of AHP and Rasch model, for the AHP, the recalculation of the received global priorities for experts in the CC (k_{AHP}) was made using the formula:

$$k_{AHP} = G_{ni}/G_{max}, \qquad (3)$$

where G_{ni} is the i-th global priority; G_{max} is the maximum value of the global priority G_{ni}.

For the comparison of the results obtained with the use of the AHP method and Rasch model, the recalculation of the total points obtained for experts in the CC (k_{MR}) was made by the formula:

$$k_{MR} = Z_i/Z_{max}, \qquad (4)$$

where Z_i is the total score for the i-th expert; Z_{max} – the maximum total score for all experts.

In the comparative analysis of the results obtained by the AHP methods and with using the Rasch model, the dispersion (variation) of the CC for experts was calculated by the formula:

$$R = k_{max} - k_{min}, \qquad (5)$$

where k_{max} is the maximum CC (equal to 1.00); k_{min} is the minimum CC obtained for a particular expert.

The values of the obtained CC for all experts in measurement of time and frequency are shown in Table 3. The indicated CC were obtained with the use of AHP (k_{AHP}), described in [2], and Rasch model (k_{RM}).

The results of the comparisons of the CC of experts in measurement of time and frequency obtained with the use of AHP and Rasch model are shown in Fig. 5.

As can be seen from Fig. 5, the comparison of CC of experts obtained with the use of AHP and Rasch model have a clear correlation. At the same time, for AHP, a larger dispersion of CC values (k_{AHP}) is characteristic: the difference between the largest and the smallest CC is R_{AHP}=0.70 (minimum CC – 0.30). For the Rasch model, the distribution of CC values (k_{RM}) is R_{MR}=0.48 (minimum CC – 0.62), i. e. less than for AHP by 31 %.

Table 3

The coefficient of competence of experts in measurement of time and frequency, obtained with the use of AHP and Rasch model

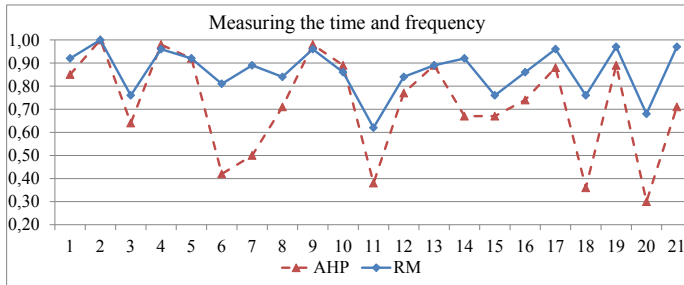| Expert	01	02	03	04	05	06	07	08	09	10	11
k_{AHP}	0.85	1.00	0.64	0.98	0.92	0.42	0.50	0.71	0.98	0.89	0.38
k_{RM}	0.92	1.00	0.76	0.96	0.92	0.81	0.89	0.84	0.96	0.86	0.62
Expert	12	13	14	15	16	17	18	19	20	21	–
k_{AHP}	0.77	0.89	0.67	0.67	0.74	0.88	0.36	0.89	0.30	0.71	–
k_{RM}	0.84	0.89	0.92	0.76	0.86	0.96	0.76	0.97	0.68	0.97	–

Fig. 5. Comparison of the competence coefficients of experts in measurement of time and frequency obtained with the use of AHP and Rasch model

The average maximum own number for technical experts obtained using AHP is 22.2. The average of the total score for technical experts, obtained with using the Rasch model is 31.9.

A general comparison of the CC of experts in metrology, obtained with the use of MAI and Rasch model, shows a clear correlation of the values obtained by experts. At the same time, a larger dispersion of CC values (R_{AHP}) is characteristic for the AHP, and for the Rasch model, the distribution of the values of the CC (R_{RM}) is lower than for the AHP.

8. Conclusions

1. The methods that are suitable for evaluating the competence of technical experts were considered in detail. The competence of technical experts (in time and frequency measurement) was evaluated according to the established criteria using the AHP and Rasch model. The results were processed using the specialized software "Competence AHP 1.1" (Ukraine) and MINISTEP 4.0.1 (USA).

2. The obtained results showed the possibility of applying the Rasch model for the analysis of the scale of expert assessment in the field of technical regulation. The analysis of the results obtained in the multidimensional Rasch model with the ECE showed that the selected scale for ECE criteria meets the requirements set by the Rasch model. The obtained measurement data for this model allow you to calculate the established statistics for both the criteria and for the evaluated experts. Only two out of twenty one (9.5 %) evaluated experts have data that are unsuitable for the analysis by the Rasch model, which indicates a low level of competence.

3. A comparative analysis of the results obtained with the use of AHP and Rasch model showed convergence, suitability and correlation of the obtained values for experts. However, the ECE using AHP, to a lesser extent, allows for the consideration of less competent experts than the ECE in the Rasch model. This is evidenced by the lower CC for AHP than the CC, obtained by the Rasch model. Thus, the AHP and the Rasch model should be used as a useful tool for comparative ECE based on objective data according to established ECE criteria for different fields of activity. At the same time, the ECE by the Rasch model allows the selection of the most competent technical experts and reject experts whose data do not correspond to a certain level of established requirements.

References

1. Velychko O. M., Gordiyenko T. B., Kolomiets L. V. Methodologies of expert's competence evaluation and group expert evaluation // Metallurgical and Mining Industry. 2015. Issue 2. P. 262–271.

2. Velychko O., Gordiyenko T. Evaluation of competence of the experts in field of metrology and instrumentations // XXI IMEKO World Congress "Measurement in research and industry". Prague, Czech Republic, 2015. 5 p.

3. Evaluation of experts competence on the measurement of electrical power using the method of analytic hierarchy / Velychko O. N., Gordienko T. B., Karpenko S. R., Kolomiets L. V. // Metallurgical and Mining Industry. 2016. Issue 11. P. 70–76.

4. Velychko O., Gordiyenko T., Kolomiets L. A comparative analysis of the assessment results of the competence of technical experts by different methods // Eastern-European Journal of Enterprise Technologies. 2017. Vol. 4, Issue 3 (88). P. 4–10. doi: 10.15587/1729-4061.2017.106825

5. Velychko O., Gordiyenko T., Kolomiets L. A comparative analysis of results of the group expert assessment of metrological assurance of measurements // Eastern-European Journal of Enterprise Technologies. 2017. Vol. 6, Issue 9 (90). P. 30–37. doi: 10.15587/1729-4061.2017.114468

6. Saaty T. L. The Hierarchon: A Dictionary of Hierarchies. Pittsburgh, Pennsylvania: RWS Publications, 1992. 510 p.

7. Saati T. L. Prinyatie resheniy pri zavisimostyah i obratnyh svyazyah: Analiticheskie seti. Moscow: Izd-vo LKI, 2008. 360 p.

8. Drake P. R. Using the Analytic Hierarchy Process in Engineering Education // International Journal of Engineering Education. 1998. Vol. 14, Issue 3. P. 191–196.

9. Chernysheva T. Yu. Ierarhicheskaya model' ocenki i otbora ekspertov // Doklady TUSUR. Upravleniya, vychislitel'naya tekhnika i informatika. 2009. Issue 1 (19). P. 168–173.

10. Kolpakova T. A. Opredelenie kompetentnosti ekspertov pri prinyatii gruppovyh resheniy // Radioelektronika, informatyka, upravlinnia. 2011. Issue 1. P. 40–43.

11. Kalinina I. O., Hozhyi O. P., Musenko H. O. Vrakhuvannia kompetentnosti ekspertiv u metodakh bahatokryterialnoho analizu v zadachakh ratsionalnoho vyboru // Nauk. pratsi Chornomor. derzh. univer. Kompiuterni tekhnolohiyi. 2013. Vol. 191, Issue 179. P. 116–123.

12. Bhushan N. Strategic Decision Making: Applying the Analytic Hierarchy Process. London: Springer-Verlag, 2004. 170 p. doi: 10.1007/b97668

13. Rasch G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institution for Educational Research, 1960. 199 p.

14. Andrich D. Rasch models for measurement. Newbury Park CA: Sage, 1988. 95 p. doi: 10.4135/9781412985598

15. Bond T. G., Fox C. M. Applying the Rasch model: fundamental measurement in the human sciences. 2nd ed. Psychology Press, 2007. 360 p. doi: 10.4324/9781410614575

16. Hagquist C., Bruce M., Gustavsson J. P. Using the Rasch model in nursing research: An introduction and illustrative example // International Journal of Nursing Studies. 2009. Vol. 46, Issue 3. P. 380–393. doi: 10.1016/j.ijnurstu.2008.10.007

17. Ehlan A. H., Kucukdeveci A. A., Tennant A. The Rasch Measurement Model // Research Issues in Physical & Rehabilitation Medicine. Pavia: Maugeri Foundation, 2010. P. 89–102.

18. Demenchenok O. G. Matematicheskie osnovy Rasch Measurement // Pedagogicheskie izmereniya. 2010. Issue 1.

19. Wright B. D., Linacre J. M. A measurement is the quantification of a specifically defined comparison. Rasch model derived from objectivity // Rasch Measurement Transactions. 1987. Vol. 1, Issue 1. P. 4–5.

20. A User's Guide to WINSTEPS®MINISTEP Rasch-Model Computer Programs. Program Manual 4.0.0 by John M. Linacre. 2017.

21. Rasch Model/Rasch Analysis: Definition, Examples // Statistics How To. URL: http://www.statisticshowto.com/rasch-model

22. Communication Validity and Rating Scales. Institute for Objective Measurement. URL: http://www.rasch.org/rmt/rmt101k.htm

23. Velychko O., Gordiyenko T. The evaluation of activity of Technical Committees of Standardization for Metrology and Measurement on national level // Journal of Physics: Conference Series. 2016. Vol. 772. P. 012007. doi: 10.1088/1742-6596/772/1/012007