A Fuzzy-Possibility Approach to a Multi-Factor Comparison of the Efficiency of the Scales for Assessing the Competencies of the Examinee

Nelya Domshenko[1] [0000-0001-8887-9205], Maria Morozova[1] [0000-0002-8514-2803], Svetlana Rubtsova[1] [0000-0003-2684-5872], Alexander Spesivtsev[2] [0000-0002-8928-4585] and Victor Lazarev[3] [0000-0002-8304-1478]

[1] Saint Petersburg State University, Universitetskaya embankment, 7–9, 199034 St. Petersburg, Russia
nelly.d@zoho.com, morozova.m@mail.ru
[2] St. Petersburg Federal Research Center of the Russian Academy of Sciences, 14th line V.O., 39, 199178 St. Petersburg, Russia
sav2050@gmail.com
[3] St. Petersburg National Research University of Information Technologies, Mechanics and Optics, Kronverkskiy prospect, 49, letter A, 197101 St. Petersburg, Russia
holod25@yandex.ru

Abstract. One of the main and essential elements of the educational process is the assessment of the quality of the test takers' competencies in the process of mastering curricula. The main final tool of this process for the teacher and the test-taker is the assigned grade according to the accepted scale. The existing commonly used and historically developed in Russia five-point system, which has been repeatedly discussed and criticized in pedagogical communities, has a fundamental drawback - it is based on an ordinal scale (points), which does not have the property of addition. The restructuring of higher education has led to the expansion of the powers of the university, in particular, the choice of the student assessment system. This made it possible to search for more convenient non-traditional scales with more differentiable estimation properties. For the first time, on the basis of a fuzzy-possibility approach, the expertise of teachers of the English language was formalized in the form of mathematical models, and then, using these models, a comparison was made of two common scales for assessing the knowledge of test-takers - the "five-point" (actually three-point) scale adopted in Russia and ECTS (European Credit Transfer and Accumulation System). It is shown that the ECTS scale is provided by a more subtle psychological and pedagogical tool for both the teacher and the student in assessing their knowledge, while the “five-point” scale does not possess such properties. The general requirements for the selected rating scales have been clarified: continuity is an inherent property of any numerical scale.

Keywords: fuzzy-possibility model, competency assessment scales, expert knowledge.

Proceedings of the 4th International Conference on Informatization of Education and E-learning Methodology: Digital Technologies in Education (IEELM-DTE 2020), Krasnoyarsk, Russia, October 6-9, 2020.

1 Introduction

Each person understands what knowledge is, although a decent definition for it has not yet been invented. The industrial stage of society's development has been replaced by a new evolutionary phase, the phase of informatization, in which the most effective and dynamic development of society is possible on the basis of the fullest possible use of available information resources, primarily knowledge, and means of their processing. Typical knowledge representation models are logical models; products; semantic networks; frames; scripts; production rules, fuzzy neural network models do not provide an opportunity to obtain analytical models, although the problem of extracting, evaluating and using knowledge is coming to the fore. But it, in the opinion of IT specialists, is "the narrowest and most complex problem", because verbal expert knowledge is difficult and inconvenient for computer processing [1].

At the same time, the teacher assessing the competence of the student is in the typical conditions of an "intellectual information and diagnostic system" [2]:. an expert is viewed as an intellectual measuring and diagnostic system. This definition applies completely to the profession of a teacher, as it is certain that they: possess intellect; carry out knowledge assessment in a specific area of a discipline; and diagnose the general extent of competence development in the examinee. This definition meets the description of expert's work in any sphere and is increasingly applied in technical and IT sciences. The decision-making mechanism in the situation of uncertainty is described with the help of models based on the expert's explicit and implicit knowledge and their professional experience [2,3].

Replacing the knowledge of an expert teacher with a mathematical model opens up ample opportunities for carrying out additional research and solving various pedagogical problems, including comprehending their preferences when assessing the quality of a student's competence. However, the optimization of the assessment of verbal linguistic competencies (knowledge) by quantitative categories depends not only on the qualifications of the teacher, but is also inextricably linked with the scales used. The need to choose effective scales for assessing competencies is most tangible in everyday teaching practice [1].

This study reflects the expertise of teachers and gives an example of constructing fuzzy-possibility models to assess students` language (linguistic) and speech competencies, and then, evaluate the effectiveness of two fundamentally different scales. The research was carried out with several groups of students studying English at B2 level (under the CEFR [4] in situations of cross-cultural foreign language communication. In order to build such models, it is essential to use a set of variables constituting the factor space within which the teacher decides whether the level of required competence was achieved [4].

2. The applied method and formulation of the problem of assessing the scales

An important condition for the study of pedagogical activity is the possibility of modeling its individual aspects, for example, to create expert systems for monitoring a student`s competence in a specific subject in the multifactorial space. [5]. At the same time, the authors themselves believe that the main reason for not entirely successful attempts to create them is the lack of a methodology for representing knowledge in a convenient form for computer processing [6].The methodology of the fuzzy-possibility approach of extracting, presenting and formalizing explicit and implicit expert knowledge for technical systems has been created [3] and is currently being tested in the field of pedagogy in teaching English at the university [4].

2.1. A Fuzzy-Possibility Approach to Building Expert Knowledge Models

Implicit expert knowledge, containing the teacher's professional experience, along with explicit ones, plays a key role in decision-making processes about the state of a learner's competence of any level of complexity. Therefore, a generalized description of the problem of extracting and formalizing explicit and implicit expert knowledge by a mathematical model is multifaceted and should contain all the necessary procedures of a systematic approach. In a set-theoretic setting, the entire procedure can be expressed by a tuple:

$$<X, \Phi_z, E_{z\,kr}, F_z, M_z, Z^E_{pr}(q,t,\mu), Z^E_{fs}(q,t,\mu), Z^E_{aa}(q,t,\mu), K, Y/\Xi>,$$

where

X – is the set of states of competencies in the English language course;

Φ_z – a set of methods of knowledge extraction;

$E_{z\,pr}$ – a set of methods of knowledge representation;

F_z – a set of methods and algorithms for formalizing knowledge;

M_z – is a set of conditions necessary for performing all operations in building models;

MZ^E (q, t, μ) – expert meta-knowledge, including explicit professional $Z^E_{pr}(q,t,\mu)$, implicit $Z^E_{pr}(q,t,\mu)$and knowledge in adjacent areas $Z^E_{aa}(q,t,\mu)$ for solving the problem of the state of competencies in the English language course;

K – is a set of quality indicators for solving problems of assessing the state of a student's competencies;

Y / Ξ – a set of classes of states of knowledge, one of which should include the result of the determination (diagnostic assessment scale).

The problem of constructing a fuzzy-possibility model is as follows. The meta-knowledge $MZ^E(q,t,\mu)$ of an expert, including professional explicit $Z^E_{pr}(q,t,\mu)$ and implicit $Z^E_{aa}(q,t,\mu)$,knowledge about the state of the student's competencies as a phenomenon needs to be transformed into a form convenient for computer processing.

The essence of the application of the fuzzy-possibility approach to the approximation of explicit and implicit knowledge of a teacher by a mathematical model is presented in Fig. 1.

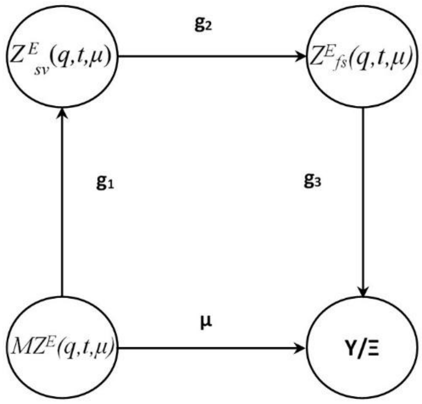

Fig. 1. A commutative diagram of the sequence of processes of extraction, presentation and formalization of explicit and implicit expert knowledge; $Z^E_{sv}(q,t,\mu)$ – set of variables, $Z^E_{fs}(q,t,\mu)$ – factor space

As follows from the analysis of Fig. 1, the expert solves the task of assessing the student's competencies directly by the ratio μ Y / Ξ, but to understand the processes occurring in their mind, this ratio has to be presented as a composition of relations

$$\mu = g_1 \circ g_2 \circ g_3,$$

where

g_1 – is the ratio of metaknowledge extraction $MZ^E(q,t,\mu)$ of an expert of a set of indicators for solving the problem of assessing the state of competencies of a particular student;

g_2 – the ratio of the representation of indicators E_{zkr} in the form of linguistic variables and the formation of factor space, in which the expert makes a decision in a particular situation;

g_3 – the ratio of formalization of expert knowledge in the form of a polynomial expression on the set of linguistic production rules of the "situation - assessment" type with diagnosing the state of the student's competencies on the scale of the factor set Y / Ξ.

Within the framework of this study, the factor space corresponding to the task of constructing a mathematical model for assessing the competencies of the test taker in English was determined by seven input variables [4].

X_1 – discourse competence when evaluating: coherence, cohesion and consistency of the examinee’s answer; and the ability to produce reasoned statements and to indulge in critical thinking. There should be a particular emphasis on the ability of students to support the statement with relevant examples, statistical data and references to the latest research in a specific area.

Grammar competence implies two factors:

X_2 – communication in a natural manner, using proper and adequate grammar structures depending on the context of the utterance.

X_3 – usage of diverse grammatical structures of the English language in accordance with the indicated level.

Lexico-semantic competence has also been considered in two aspects:

X_4 – skills to select vocabulary depending on the situation, diversity of vocabulary and, as a result, absence of repetition.

X_5 – adequacy of lexical elements.

X_6 – phonetic competence, which implies proper intonation, speech fluency, presence or absence of pronounced accent, pausing or mispronunciation of separate words.

X_7 – social interaction competence. This has been limited to the ability of the examinee: to understand direct or implied sense (meaning) from the speech of the interlocutor; as well as to use the language for specific purposes depending on the characteristics of social and professional interaction including the situation and status of the interlocutor-examiner.

The choice of competences corresponds to: the systemic approach to the examinee knowledge assessment procedure; and the overall index of course competence Y.

The main steps of building a fuzzy-possibility model for the assessment of the examinee`s competencies using the ECTS scale were discussed [4]. Moreover, each of the above competences, including Y in the factor space, was presented as a linguistic variable (Figure 2).

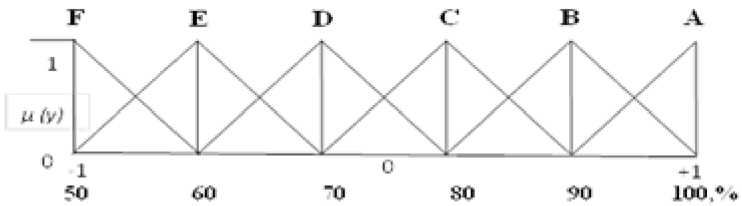

Fig. 2. Y as a linguistic variable

The linguistic variable (Fig. 1) is used for converting a verbal grade definition into quantitative information. It contains three scales on the x-axis: linguistic linguistic F, ..., A (the upper scale); numerical (to convert into natural scale) from 50 to 100%, respectively; and independent variables, a standardized scale to be used in experimental design theory [«-1» – lowest grade E; «+1» – highest grade A]. On the y-axis, there is the membership function scale of μ (y) grade. This means that the closer the grade is to the class mode, the higher is its accuracy.

The converted seven-point ECTS (European Credit Transfer and Accumulation System) grading scale, which is shown in Table 1, and the generally accepted three-point scale in Russia as a part of the “five-point”, were chosen as the scales for assessing the degree of mastering both individual competencies and the subject as a whole.

Table 1. Examinee grading scale

Grade	**Percentile %**	**Verbal definition**	**Corre-sponding score**
A	90-100	Excellent	5
B	80-89	Very good	4+
C	70-79	Good	4
D	60-69	Satisfactory	3
E	50-59	Performance meets the minimum criteria	3-
FX		Unsatisfactory	2
F		Fail	1

Table 2 shows a fragment of the examinational matrix with the expert's grades in linguistic form on the ECTS scale, where each line stands for an implicative condition-action rule “if …, then …” (“situation” – “grade”).

Table 2. Excerpt from the examinational table with the experts’ answers in linguistic form

	X1 – discourse competence	**X2 – communication in a natural manner**	**X3 – usage of diverse grammatical structures**	**X4 - skills to select vocabulary depending on the situation**	**X5 - adequacy of lexical elements**	**X6 – phonetic competence**	**X7 – social interaction competence**	**Y – overall index of competence**
	x1	x2	x3	x4	x5	x6	x7	Y
1	-1	-1	-1	-1	-1	-1	1	E
2	1	-1	-1	-1	-1	-1	-1	E-D
3	-1	1	-1	-1	-1	-1	-1	E
4	1	1	-1	-1	-1	-1	1	C
…	…	…	…	…	…	…	…	…

62	1	-1	1	1	1	1	-1	B-A
63	-1	1	1	1	1	1	-1	B
64	1	1	1	1	1	1	1	A

Conversion of verbal expert grades into numerical form using the scales of Figure 2 and processing these data using the methods of experimental design theory [2,3] has resulted in the following analytical equation:

$$Y = 74.80 + 6.83x_1 + 5.08x_2 + 5.07x_3 + 3.90x_4 + 4.49x_5 + 3.32x_7 - 1.36x_2x_3x_5 + 1.37x_2x_5x_6 + 1.95x_3x_4x_6, \qquad (1)$$

where only significant coefficients are given, and all variables are presented on a standardized scale according to the formulas:

$x_i = \frac{X_i - \bar{X_i}}{\Delta X_i}$, $\bar{X}_i = \frac{X_{max} + X_{min}}{2}$, $\Delta X_i = \frac{X_{\max} - X_{min}}{2}$, i - number of variables.

The constructed model (1) now makes it possible to conduct a study to assess the effectiveness of the scales. As grading scales for the learning outcomes the ECTS scales and the five-point scale adopted in Russia were selected.

2.2. Comparison of the effectiveness of the scales

The research methodology was as follows. The competence of one and the same group of students was simultaneously assessed on both scales by the same teacher. This approach makes it possible to assess only the effectiveness of the scales. The criterion for the scale's effectiveness was the degree of consistency between the expert-teacher`s assessments and the constructed fuzzy-possibility model calculations (1) for assessing course competencies based on the selected factor space.

The adequacy of analytic equation (1) was verified by two criteria: correlation between the expert assessment and calculations based on the values in (1) (Figure 3 a) as confidence that the calculations according to (1) correspond to the expert`s opinion on the problem under study as a whole, and the correlation between calculated values in (1) and grades awarded to a group of eight students by the teachers who were not familiar with this methodology (Fig. 3 b).

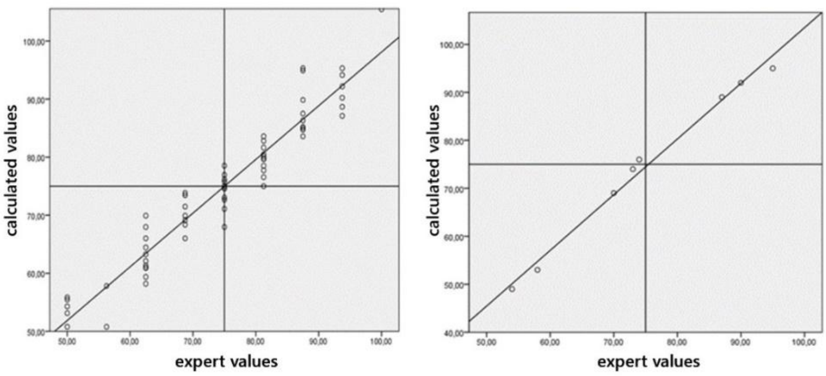

Fig. 3. Estimation of the adequacy of calculations according to (1): a - knowledge and experience of experts; b - data from an independent experiment

The analysis of the scattering of points around the theoretical regression line (the diagonal of the square) indicates the expert's high professionalism (Figure 3a, the correlation coefficient R = 0.98) and the adequacy of the calculations according to the model (1) in their opinion. At the same time, Figure 3b indicates, firstly, the qualified assessment of the competences of the examinee by an independent teacher, and, secondly, the effectiveness of the selected assessment scale throughout the course as a whole (correlation coefficient R = 0.96).

Studies of the generally accepted in Russia five-point (and in practice, "three-point" scale 3, 4 and 5) using the same methodology and the same factor space led to the mathematical expression:

$$Y = 4.01 + 0.262x_1 + 0.207x_2 + 0.207x_3 + 0.152x_4 + 0.176x_5 + 0.121x_7 - 0.066x_3x_6 - 0.059x_1x_2x_4 - 0.058x_2x_3x_5 + 0.098x_3x_4x_6. \quad (2)$$

The graphs similar to Figure 3 on a “three-point” scale, as shown in Figure 4, allow us to conclude that the teacher's competence in building the model (Figure 4a) “for some reason” decreased - the correlation coefficient became R = 0.925, and the effectiveness of the selected scales of this type, when evaluating the examinees` learning outcomes for the subject as a whole by the same teacher (Figure 4b), decreased quite significantly - the correlation coefficient R = 0.74.

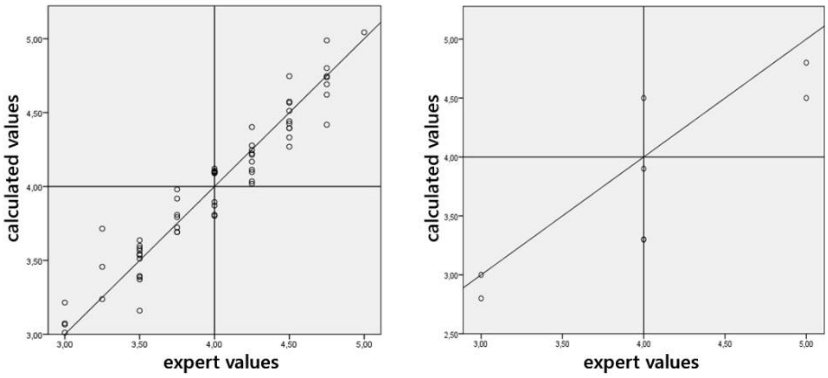

Fig. 4. Evaluating the adequacy of calculations based on (2): a - knowledge and experience of experts; b - independent experiment data

A noteworthy fact is that the nonlinear terms, in models (1) and (2), do not coincide either in composition or in number, despite the identical verbal assessments of situations reflecting the teacher`s expertise on the problem under consideration.

The conducted studies allow us to draw a conclusion which is not in favor of the Russian scaling system for assessing the student`s competencies.

3. Discussion of results

The most important advantage of the ECTS scale is that it is both numeric and interval. This makes it possible to apply all mathematical statistics methods to the test scores, for example, adding together, calculating arithmetic means, etc.

The “five-point” (actually three-point) scale has low efficiency for fundamental reasons. There is a void between the quality points on the scale. The point scale is ordinal, in which, by definition, there is no addition operation. Points are not numbers, although they can be, as in this case, indicated by numbers.

The study to assess the effectiveness of two fundamentally different scales became possible due to the use of a fuzzy-possibility approach. The building of mathematical models has clearly demonstrated the possibility of conducting special studies with obtaining visual, qualitatively new results. The conclusion about the advantage of the ECTS interval scale over the point scale is not entirely new, but this study provided clear evidence of the low efficiency of ordinal scales for assessing students` competencies using a rigorous mathematical apparatus.

It should be noted the only drawback of the ECTS scale is that it is discontinuous at the points of contact of modal intervals, for example A and B (89-90). At such points, it is advisable that the intervals partially intersect, as is the case in the fuzzy-possibility approach (see Fig. 2). This would give the teacher a psychological advantage in justifying "why 89 over 90".

4. Conclusion

The fuzzy-possibility approach of building mathematical models based on explicit and implicit expert knowledge is provided by a significant tool that allows you to successfully solve problems of a methodological and applied nature in pedagogical sciences, as in this case - assessing the effectiveness of the examinee`s competences on various scales. At the same time, it is shown that the interval European scale ECTS (European Credit Transfer and Accumulation System) allows teachers to assess the knowledge of students more differentially and explain the given grade more reasonably. The five-point scale adopted in Russia is ordinal and does not possess such qualities, which makes it less effective in comparison with ECTS.

As is shown in the study, fuzzy-possibility models also make it possible to preserve and replicate the experience of highly qualified experts, for example, for the practical use of junior teachers. They can be applied to conducting in-depth mathematical research, as in this example, or using as a knowledge base for building expert systems of any complexity.

In addition, it can be stated that the typical models of knowledge representation, mentioned in the introduction, were supplemented by fuzzy-possibility models in the form of analytical expressions.

References

1. Gavrilova T.A., Khoroshevsky V.F.: Knowledge base of intelligent systems skills [in Russian]. PITER, 382 (2000).
2. Spesivtsev A.V., Domshenko N.G .: Expert as an "intelligent measuring and diagnostic system" skills [in Russian]. Sat. reports. XIII International Conference on Soft Computing and SCM Measurements, 23-25 July 2010, St. Petersburg, Vol.2. pp.28-34 (2010).
3. Ignatiev MB, Marley VE, Mikhailov VV, Spesivtsev AV: Modeling weakly formalized systems based on explicit and implicit expert knowledge skills [in Russian]. SPb: POLYTECH-EXPRESS, 430 (2018).
4. Domshenko N.G., Morozova M.N., Rubtsova S.Yu., Spesivtsev A.V. : Assessment of the test taker's competence based on logical-linguistic skills [in Russian]. International scientific research journal INTERNATIONAL RESEARCH JOURNAL, Yekaterinburg, 2 (80), pp. 138-142. (2019)
5. Chvanova MS, Kiseleva IA, Molchanov A.A.: Problems of using expert systems in education skills [in Russian]. Bulletin of the Tambov University. Series: Humanities. No. 3 (119) pp.39-47 (2013).
6. Kotova E.E., Stepanov A.G.: Statement of the problem of assessing the competence of a graduate by methods of soft measurements skills [in Russian]. XXI International Conference on Soft Computing and SCM Measurements 23-25 May 2018 .-- St. Petersburg, Vol. 2 - P.281-282 (2018).