CRITERIA AND DIAGNOSTIC TOOLS FOR ASSESSING THE EFFECTIVENESS OF VISUALIZATION IN ENGINEERING EDUCATION

Yuldasheva Muhayyo Alimbayevna - independent researcher at Urgench State University.

Annotation: The article addresses the problem of objectively assessing the effectiveness of visualization tools in engineering education. A system of four criteria is substantiated: cognitive, operational-activity, motivational-value and reflective, each specified through indicators and three levels of development. A set of diagnostic instruments is described, including concept-map scoring, spatial ability tests in the spirit of S. Sorby's research, an analytic rubric for student-made diagrams and questionnaires. The four-level evaluation model of D. Kirkpatrick is adapted to the evaluation of visualization-based teaching methods. The design of a pedagogical experiment with control and experimental groups and procedures for the statistical verification of results is outlined, without anticipating its outcomes.

Keywords: effectiveness assessment, criteria and indicators, diagnostic tools, concept-map scoring, spatial abilities, rubric, Kirkpatrick model, pedagogical experiment.

The introduction of visualization tools into the teaching of technical disciplines is now widely supported by teachers, and many departments actively use animations, interactive models, concept maps and simulations. However, the question of how to establish whether these tools actually improve learning is often left unanswered. Teachers tend to rely on impressions, which may be misleading. The Law of the Republic of Uzbekistan "On Education" names the quality of education among the basic principles of state policy in this sphere [1], and quality cannot be ensured without reliable means of measuring results. Therefore, any methodology for using visualization in engineering education must include a well-founded system of criteria,

indicators and diagnostic tools that make it possible to assess its effectiveness.

In pedagogical research, a criterion is understood as a characteristic on the basis of which the state of a phenomenon is evaluated, while indicators are observable and measurable manifestations of the criterion. Levels describe qualitatively distinct stages of development of the measured quality. In the research tradition devoted to the professional competence of vocational and engineering educators, this triad of criterion, indicator and level is used as the main instrument of diagnosis [2]. In relation to visualization, it is important to recognise that its effect is not limited to knowledge acquisition; it also concerns the ability to work with visual representations, the attitude towards them and the awareness of one's own ways of thinking. Consequently, a multi-criterion approach is required.

The first criterion, cognitive, characterises the completeness, systematicity and depth of the student's knowledge of the subject content mediated by visualization. Its indicators include knowledge of key concepts and laws, understanding of the relationships between them, the ability to interpret graphs, diagrams and schemes, and the ability to move between different forms of representation, for example from a circuit diagram to a phasor diagram or from a beam scheme to a bending moment diagram. The second criterion, operational-activity, reflects the ability to use and create visual representations in solving engineering problems. Its indicators are the correct construction of free-body diagrams, kinematic schemes and drawings, the use of diagrams as a means of analysis rather than illustration, and the development of spatial abilities, such as mentally rotating an object or reconstructing a three-dimensional part from its projections. The third criterion, motivational-value, characterises the student's interest in the subject, attitude towards visual methods of learning and the recognition of the professional value of graphical literacy for an engineer. The fourth criterion, reflective, covers the student's ability to evaluate the

quality of his or her own and others' visual representations, to identify errors in them and to choose consciously the type of representation appropriate to a task.

For each criterion, three levels can be distinguished. At the low, or reproductive, level the student recognises standard visual representations and reproduces them with the support of a model, but does not use them independently to solve problems and does not see their connection with theoretical content. At the intermediate, or productive, level the student independently constructs typical diagrams and schemes, interprets them correctly and uses them in solving standard problems, while experiencing difficulties in non-standard situations. At the high, or creative, level the student freely translates information between different forms of representation, constructs visual models for new problems, evaluates their adequacy critically and explains the choice of representation. The levels should be concretised for each discipline.

The cognitive criterion can be diagnosed with the help of concept-map scoring. Students are asked to construct a concept map on a topic, such as "Stress and strain" in strength of materials, from a given list of concepts or freely. Novak and Gowin proposed a scoring scheme that awards points for valid propositions, hierarchical levels, cross-links and examples, with cross-links weighted more heavily because they reflect integration of knowledge [3]. An alternative approach is the comparison of the student's map with a criterion map constructed by an expert teacher, in which the proportion of correct propositions is determined. Concept-map scoring is valuable because it reveals the structure of knowledge, but it requires preliminary training of students in map construction and agreement between assessors, which should be checked by independent scoring of a sample of maps by two experts.

The operational-activity criterion, as far as spatial abilities are concerned, can be assessed using standardized spatial tests. S. Sorby's long-term research on engineering students showed that spatial skills, particularly the ability to mentally rotate three-

dimensional objects, can be measured with established psychometric instruments and developed through targeted instruction, and that their improvement was associated with better performance and retention in engineering programmes [4]. Following this approach, a mental rotation test or a test requiring the identification of an object from its orthographic projections can be administered before and after a course of engineering graphics in which visualization tools are used. In addition, discipline-specific tasks can be developed, for example identifying the correct axonometric view of a part or determining the shape of a section, which link spatial ability directly to course content.

The ability to create visual representations requires a different instrument, namely an analytic rubric for evaluating student-made diagrams. Such a rubric may include the following dimensions: correctness of content, meaning that all forces, elements or quantities are represented and none are fabricated; compliance with conventions and standards, including symbols, line types, sign rules and dimensioning; structural clarity, that is, the logical arrangement of elements and the absence of clutter; the integration of labels and explanations; and the functionality of the diagram, meaning the extent to which it actually supports the solution of the problem. Each dimension is described at three levels corresponding to the levels described above, and the rubric is communicated to students in advance, so that it also serves as a tool for formative assessment and self-assessment.

The motivational-value and reflective criteria are diagnosed mainly by questionnaires and self-assessment instruments. A questionnaire with a Likert-type scale can include statements on interest in the discipline, the perceived usefulness of visual materials, confidence in one's own ability to construct diagrams and the importance of graphical literacy for future professional activity. Learning style instruments in the tradition of Felder and Silverman, who distinguished, among other dimensions, between visual and verbal learners in engineering education, can be used

to take into account individual differences when interpreting the results [5]. The reflective criterion can be assessed through reflective written tasks in which the student compares two versions of a diagram, explains which one is more appropriate and why, or describes how a visual representation helped him or her to solve a problem.

To organise these instruments into a coherent evaluation system, it is useful to adapt the four-level model of D. Kirkpatrick, originally developed for the evaluation of training programmes [6]. At the first level, reaction, the evaluation determines how students perceive the visualization tools: whether they find them clear, useful and engaging. Questionnaires administered after the use of a new visual module serve this purpose. At the second level, learning, the increase in knowledge and skills is measured by means of tests, concept-map scoring, spatial tests and rubric-based evaluation of diagrams. At the third level, behaviour, the question is whether students transfer the acquired skills to other contexts: whether they spontaneously construct diagrams when solving problems in subsequent courses, in course projects and in internships. At the fourth level, results, long-term outcomes are considered, such as academic performance in related disciplines, the quality of final qualification works and the assessments of employers. The adapted model emphasises that a positive reaction of students, which is the easiest to obtain, does not in itself prove the effectiveness of visualization.

The final verification of the effectiveness of a visualization-based methodology requires a pedagogical experiment. Its design involves the selection of control and experimental groups from students of the same speciality and year of study, taught by the same or equally qualified teachers under the same curriculum. In the experimental groups, teaching is conducted according to the proposed methodology using the developed visualization tools, while in the control groups the traditional teaching of the same content continues. Before the experiment, both groups are assessed with the

instruments described above to establish the equivalence of their initial levels; after the formative stage, the same or parallel instruments are used again. Since random assignment of individual students is rarely possible in a university, the design is usually quasi-experimental, which requires particular care in establishing initial equivalence and in describing possible confounding factors.

The statistical verification of results should correspond to the nature of the data. For level distributions measured on an ordinal scale, nonparametric criteria such as Pearson's chi-square test or the Mann–Whitney U test are appropriate for comparing control and experimental groups. For test scores that can be treated as interval data and are approximately normally distributed, Student's t-test can be applied. Beyond the determination of statistical significance, it is advisable to calculate an effect size, which characterises the magnitude of the difference independently of the sample size. Hattie's synthesis of meta-analyses showed the value of interpreting effect sizes against the typical effect of educational interventions, rather than considering any positive difference as a success [7]. This position is particularly relevant to visualization, since meta-analytic research indicates that its effects vary with the type of representation and the kind of knowledge being learned; for example, Höffler and Leutner found that instructional animations were on average more effective than static pictures, especially when the animation was representational rather than decorational and when procedural-motor knowledge was involved [8]. The experiment should therefore be designed so that the effects can be analysed separately for different criteria and types of tools.

In conclusion, the assessment of the effectiveness of visualization in engineering education requires a multi-criterion approach that covers cognitive, operational-activity, motivational-value and reflective components. The proposed system of criteria, indicators and levels, combined with concept-map scoring, spatial tests, a rubric for student-made diagrams and questionnaires, and organised according to the

adapted Kirkpatrick model, provides a basis for a pedagogical experiment whose outcomes can be verified statistically. The application of this system will make it possible to move from intuitive judgements about the benefits of visualization to evidence-based conclusions.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. O‘zbekiston Respublikasining “Ta’lim to‘g‘risida”gi Qonuni. 2020-yil 23-sentabr, O‘RQ-637-son.

2. Muslimov N.A. va boshq. Kasb ta’limi o‘qituvchisining kasbiy kompetentligini shakllantirish texnologiyasi. – Toshkent: Fan va texnologiya, 2013.

3. Novak J.D., Gowin D.B. Learning How to Learn. – Cambridge: Cambridge University Press, 1984.

4. Sorby S.A. Educational research in developing 3-D spatial skills for engineering students // International Journal of Science Education. – 2009. – Vol. 31(3). – P. 459–480.

5. Felder R.M., Silverman L.K. Learning and teaching styles in engineering education // Engineering Education. – 1988. – Vol. 78(7). – P. 674–681.

6. Kirkpatrick D.L., Kirkpatrick J.D. Evaluating Training Programs: The Four Levels. 3rd ed. – San Francisco: Berrett-Koehler, 2006.

7. Hattie J. Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement. – London: Routledge, 2009.

8. Höffler T.N., Leutner D. Instructional animation versus static pictures: A meta-analysis // Learning and Instruction. – 2007. – Vol. 17(6). – P. 722–738.