Natural-scientific cases as an instrument for assessment of logical thinking

M V Solodikhina¹,², N I Odintsova¹ and E E Odintsova²

¹Moscow Pedagogical State University (MPGU), 1/1 M. Pirogovskaya St., Moscow, 119991, Russian Federation
²Peoples' Friendship University of Russia (RUDN University), 6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

E-mail: solmari@inbox.ru

Abstract. Logic is the basis of both critical thinking and the scientific method of nature cognition, so it can be effectively developed when studying natural science disciplines. The existing tools for assessing logical thinking require a lot of time to work with materials that are barely related to natural science. Instead, it is proposed to use natural science cases, which can serve simultaneously as both a development tool and a diagnostic tool for logical thinking. The aim of the study was to test the criterion validity of such cases for assessing logical thinking. The study involved students of the Master’s program “Modern Natural Science” of the Moscow Pedagogical State University and the bachelor’s program of the Peoples’ Friendship University of Russia. The diagnostic results showed a high degree of correlation between the results obtained by the subjects when solving cases with the results of four independent assessments: Test of Logical Thinking, the Group Assessment of Logical Thinking, solving a set of logical problems and expert assessment of students’ work on a natural science project.

1. Introduction

Development of critical thinking has become one of the main strategies of modern education [1]. Critical thinking is based on logic. In many concepts critical thinking is practically equated with informal logic [2]. Logic plays a central role in scientific cognition and natural science in general [3], so the emphasis on the formation of logical thinking in teaching natural sciences is naturally determined. To track the impact of training on the development of logical thinking in students that study natural sciences, it is necessary to conduct regular diagnostics. Most often, in such cases, researchers (for example, [4,5]) use the Group Assessment of Logical Thinking (GALT) and Test of Logical Thinking (TOLT) tests, in which students select both the answer and its rationale from a number of alternatives. However, the test form for the assessment of logical thinking contradicts its essence in many ways: in logically oriented concepts of critical thinking, skills related to argumentation [6] lie at the core, and in tests students do not formulate a justification for their answer, but choose from the proposed range. One could resort to typical tasks for logic (sophisms, paradoxes, etc.), but they do not have natural scientific content. In this case, diagnostics takes time away from studying the natural science materials provided in the program. The third way is the solution of typical natural-scientific exercises aimed at finding the only correct answer when all the necessary and sufficient data are available. When solving such problems, the student often relies on the use of standard schemes and does not give reasons for the choice of the solution.
method and the reliability of the answer. This method provides experts with a minimum base for assessing the formation of students’ logical thinking.

We suggest using educational natural-scientific cases based on real problem situations to assess the logical component of critical thinking. Most of the case tasks require to perform such logical operations as analysis, interpretation, evaluation and construction of argumentation, which are the main procedures of informal logic [7]. This article presents the results of an experiment carried out using a set of cases on integrated natural-scientific disciplines developed by M.V. Solodikhina [8]. The substantive validity of these cases has been proven earlier [9] using the experts’ assessment method.

The purpose of the experiment was to determine the criterion validity of tasks of natural-scientific cases to assess the level of students’ logical thinking: to what extent the results of performing certain tasks of natural-scientific cases and the results of assessing logical thinking using the aforementioned methods correlate with each other.

The next section describes a typical natural-scientific case and provides a quick overview of one of them. It is followed by a description of the experiment, the principles of the selection of tasks to test logical thinking, as well as the methodology for assessing the criterion validity of the selected tasks. The section on the results and their interpretation compares the calculated validity coefficient of the logical tasks of cases with four external independent criteria. The role of case studies as a tool for the formation and assessment of logical thinking is emphasized in the conclusion.

2. An example of a natural-scientific case
Each case consists of three parts: a historical or topical problem (or a chain of logically related problem situations); tasks with an open answer, which help to understand these problem situations and connect them with each other (about 30 tasks); and reference materials.

Let us consider the case titled “Mendeleev” [8] as an example, in which a chain of problem situations allows students to understand how scientists overcome difficulties on their way to discoveries.

The first problem situation is related with the systematization of 63 chemical elements known by the middle of the 19th century in accordance with the principles proposed by D. Mendeleev (figure 1). To begin with, students are invited to arrange the cards with chemical elements (figure 2) in a row in accordance with their 1) atomic masses, 2) chemical properties, 3) higher oxides. The latter are proposed to be compared with a fragment of one of the first periodic tables (PT) (table 1). Finding the “out-of-the-box” value requires logic and proposing a way to solve the problem requires creativity. One can check whether their solution coincides with the one proposed by Mendeleev by comparing it with one of the first versions of the periodic table (figure 3).

![Figure 1](image1.png)
**Figure 1.** The periodic table (PT) construction principle.

![Figure 2](image2.png)
**Figure 2.** Cards indicating: symbol; atomic mass known by 1869; higher oxide; some properties of the element.
Table 1. Fragment of one of the first PT with the indication of higher oxides for each group.

| Group | I   | II  | III | IV  | V   | VI  | VII | VIII |
|-------|-----|-----|-----|-----|-----|-----|-----|------|
|       | R^2O | R^2O^2 / RO | R^2O^3 | R^2O^4 / RO^2 | R^2O^5 | R^2O^6 / RO^3 | R^2O^7 | R^2O^8 / RO^4 |

The following problem situation is caused by the identification of at least three inconsistencies in the arrangement of elements with the principle of systematization proposed by D.I. Mendeleev (figure 1). Mendeleev knew about some of them, which can be assumed based on the question marks in the table (figure 3). To propose a rational hypothesis explaining the anomalies, students should revise their knowledge and come to the conclusion that information about the structure of the atom is needed to clarify the principles of the PT construction. Subsequent tasks make students follow the historical way of clarifying the models of the atom (by Thomson, Rutherford, Bohr, Schrödinger) with an analysis of their advantages and disadvantages.

![Figure 3](image_url)

Figure 3. One of the first PT options. The dark ovals show the atomic masses of the known elements corrected by D.I. Mendeleev, and the light ovals show the atomic masses of the elements not yet discovered by the time the PT was compiled, which were predicted by D.I. Mendeleev.

Discussion of experiments aimed at the determination of the structure of the atom leads to the following problem situation: student of E. Rutherford, D. Hevesy, was unable to separate lead (207Pb) from radium-D (modern name: 210Pb), which had the same chemical properties. Students are invited to conduct their own reasoning according to the following algorithm: 1) arrange these elements in the PT in accordance with the Mendeleev systematization principle; 2) compare the hypothesis of Mendeleev with the hypothesis of N. Bohr, to whom D. Hevesy turned to: if there are chemically indistinguishable elements of different atomic weights, differences in the chemical properties of atoms do not depend on differences in weight; 3) build a logical chain leading to the conclusion on which characteristic of the atom determines its chemical properties, taking into account the experimental fact established by H. Moseley (another student of E. Rutherford) that the chemical properties of an element depend on its ordinal number in the PT.

A chain of tasks leads to the final problem situation of the case connected with the clarification of the origin of chemical elements that make up the human body. Calculations in Excel make it possible to estimate the volume and mass of all chemical elements of the human body and compare them with the volume and mass occupied by cells, molecules, atoms and elementary particles [10].

3. Research method

The sample of the main study consisted of 52 students of the Master’s program “Modern Natural Science” of the Moscow Pedagogical State University (MPGU) of six subsequent years of enrollment.
from 2014 to 2019. Each student was supervised for two semesters while studying three different disciplines and completing a science project.

As a key to critical thinking many researchers define the ability to interpret and analyze information, draw logically correct and acceptable conclusions (inferences), and test the validity and acceptability of judgments, arguments, beliefs and interpretations (evaluation) (for example, [2]). In accordance with this, tasks of three types were selected in each of the cases: 1) interpretation and analysis of information; 2) formulation of conclusions and their evaluation; 3) construction of an algorithm. All tasks were rated on a 10-point scale.

The tasks of the first type are usually formulated as “continue the chain”, “insert the missing”, “find an analogy”, etc. For example, it is proposed to construct consistent chains of conclusions based on the postulate that moving charges continuously emit electromagnetic waves: “an electron moving in an orbit ... waves, therefore, a) its emission spectrum is ..., b) the energy of the electron ..., the radius of the electron’s orbit ..., that is, the movement occurs along ... (indicate the shape of the trajectory), as a result of which the electron ... (indicate what will eventually happen to the electron)”[8]. The known facts about the emission spectra and the lifetime of atoms serve as verification. Also, the first type includes tasks in the form of logical “puzzles” that challenge the ability to detect analogies: for example, fill in tables 2 and 3 using figure 3 and modern PT. An example of a logical task for finding inconsistency is spotting “anomalies” in the PT (figure 4).

Table 2. Atomic masses of known elements “corrected” by Mendeleev.

| Name of the element | Atomic mass | Properties                     |
|---------------------|-------------|--------------------------------|
|                     | Established by the middle of 19th century | Calculated by Mendeleev | Modern |
|                     |             |                               |        |
| 75.6                | Softest metal, chemically similar to Ti   |                   |        |
| 123.4               | Alkali metal                                     |                   |        |
| 60                  | Metal, oxide R₂O₃                                  |                   |        |
| 94                  | Metal, oxide R₂O₃                                  |                   |        |

Table 3. Elements “predicted” by Mendeleev and the values of their atomic masses.

| The name given by Mendeleev | Modern name | Atomic mass (figure 3) | Modern atomic mass |
|-----------------------------|-------------|------------------------|--------------------|
| eka-aluminium               | Modern name | Atomic mass (figure 3) | Modern atomic mass |
| eka-silicon                 | Modern name | Atomic mass (figure 3) | Modern atomic mass |
| eka-manganese               | Modern name | Atomic mass (figure 3) | Modern atomic mass |
| dvi-manganese               | Modern name | Atomic mass (figure 3) | Modern atomic mass |
| eka-cesium                  | Modern name | Atomic mass (figure 3) | Modern atomic mass |

Figure 4. Fragments of modern PT.

The tasks of the second type require to find and analyze the reliability of the arguments, make conclusions based on them that are substantiated deductively or inductively. For example, it is proposed to make a reasoned assumption why predictions of Mendeleev for elements with atomic masses of 0.17; 0.4; 54; 140; 146; 155; 170; 175 did not come true, and why the confirmed predictions proved the predictive power of the PT, and the unsuccessful predictions had no negative consequences for chemistry. If a student only presents a conclusion, he or she receives 10% of the maximum score, if, in addition to the conclusion, the facts underlying the conclusion are given, but there is no argumentation - up to 40%. A student can receive the maximum score for argumentation with links to reliable sources. The importance of teaching students to justify their choice of answer is also related to the fact that if students find any acceptable answer, they stop considering alternatives [11].
The third type of tasks relies on the ability of logical thinking to create algorithms. For example, students are asked to recreate possible algorithms of actions by G. Hevesy and A.P. Vinogradov. G. Hevesy discovered that radium-D emits electrons that can be detected, and used his discovery - the tagged atom method - to test the hypothesis that the canteen used leftovers from the previous meal in the preparation of food (and he was right). A.P. Vinogradov used the oxygen isotope $^{18}\text{O}$ to find out where from $\text{O}_2$ is released during photosynthesis: from $\text{H}_2\text{O}$ or $\text{CO}_2$. Figure 5 acts as a reference material, where the color of the oxygen symbols in the formulas is a hint.

The criterion validity of the logical tasks of cases was calculated according to the formula proposed in [12] for calculating the validity coefficient:

$$V = \frac{\sum_{i=1}^{n}(Y_i \cdot \bar{y}) - \overline{Y} \cdot \bar{y}}{\sqrt{\sum_{i=1}^{n}Y_{i}^{2}}} \cdot \frac{n}{n-1},$$

Where the numerical sequences $y_1, y_2, ..., y_n$ and $Y_1, Y_2, ..., Y_n$ are the assessments of knowledge of $n$ subjects ($n = 52$), obtained, respectively, using the external criterion and case tasks; $\bar{y}$ and $\overline{Y}$ is their arithmetic mean; $S_y$ and $S_Y$ are their standard deviations. In the cases, each of the $Y_i$ values was calculated as the subject’s average score for all 132 logical tasks of all solved cases. In accordance with [12] the validity is considered low if $V \in [0; 0.3)$, the validity is considered medium if $V \in [0.3; 0.6)$, and if $V \in [0.6; 1]$ – the validity is considered high.

The coefficient of validity of case logic tasks was calculated for four independent assessments of logical thinking. The first two external assessment criteria were the TOLT and GALT tests in Russian with the confirmed validity [13][14]. TOLT consists of 10 questions designed to measure five modes of formal reasoning: controlling variables, proportional, combinatorial, probabilistic and correlational reasoning. GALT contains 12 question instruments that tests six logical operations: Conservation, Controlling Variables, Proportional, Combinatorial, Probabilistic and Correlation Reasoning. Both tests had tasks that suggested to justify the choice of an answer by choosing from a number of alternatives. The third criterion was a set of logical problems like Einstein’s riddle about 5 houses, “find an extra figure in a row”, “continue a logical sequence”, tasks on the “Sherlock Holmes method” [15], sophisms, historical logical paradoxes, etc. (28 tasks in total). The fourth criterion is an expert assessment of students’ work on an individual natural science project according to such parameters as the quality of justifying the relevance of the project; accuracy in formulating the topic, purpose, object and subject of the research; logical consistency of purpose, hypothesis and objectives.

The tool proposed for diagnostics (natural-scientific cases) takes up a large amount of study time, which makes it irrational to use it solely for diagnostic purposes, and the use of problem-situational tasks (cases) allows students to nourish critical thinking skills directly when studying academic disciplines [16]. Therefore, a study was carried out to assess its impact on the dynamics of changes in the level of logical thinking. The sample of the study consisted of 32 undergraduate students of the Faculty of Physics, Mathematics and Natural Sciences of the Peoples’ Friendship University of Russia (RUDN University). Every week, students completed one of the natural science cases corresponding to the content of one or another session of the course “Concepts of modern natural science”.

Figure 5. Diagram of the photosynthesis process.
Some of the tasks of the cases performed a teaching function, and some of the specially selected tasks from the same cases (a total of 76 tasks of the three types described above out of 8 cases) served as an indicator of changes in the level of students’ logical thinking.

4. Results and their interpretation

The analysis (table 4) showed a high correlation of the diagnostic results based on natural-scientific cases with the results of four independent tools for assessing logical thinking (V>0.6). The highest correlation coefficient (0.87) was obtained for the expert assessment, which included the most complete coverage of all possible logical operations in the work of students.

Table 4. Numerical coefficients of validity of the logical tasks of cases.

| Independent assessment tools | TOLT | GALT | A set of logic problems | Experts |
|-----------------------------|------|------|-------------------------|---------|
| Criterion of validity       | 0.82 | 0.61 | 0.73                    | 0.87    |

The correlation of assessment methods for each of the students who took part in the experiment is shown in figure 6. The height of the bars corresponds to the average scores obtained by each of the students: a) when completing science cases (black); b) when solving TOLT, GALT and a set of historical logic problems (shaded); c) given by experts (gray). The obtained data give grounds to assert that natural science cases can serve as a tool for diagnosing the logical thinking of students.

Figure 6. Comparison of the results of assessing the logical thinking of 52 students using three different methods. The vertical axis is the student’s score on a 10-point scale, the horizontal axis is the student’s ordinal number.

The results of the assessment study reflect the dynamics of changes in the logical thinking of students obtained using our proposed tool. Figure 7 shows that students initially coped well with the tasks of the first type, and the score barely increased throughout the study. Students of the second type initially either skipped or wrote down only the conclusion without argumentation, but by the fourth case problem they began to argue all tasks of this type. On the interval [1;4), the increase in score was ensured due to the increase in the number of reasoned answers, and starting from case 4 - due to the improvement in the quality of argumentation. The situation is similar with tasks of the third type: the initial low average score is associated with the failure of some students to complete this type of task or extremely schematic description of algorithms. After the fourth case, all students began to complete the tasks, and the increase in scores was due to the more detailed algorithms.

The line in figure 7 connects the points corresponding to the average score for all three types of tasks. It can be assumed that the high growth of scores in the interval [1;4) is explained by students getting used to such tasks, and the growth in the interval (4;8] reflects change in the level of their logical thinking. Therefore, the calculation of the empirical value of the t-test for dependent samples, where the initial data had a normal distribution, was carried out based on the average difference in values between the results of the fourth and eighth cases (table 5). The level of significance allows us to talk about the positive impact of cases on the development of logical thinking of students. It should be noted that a significant increase in scores was observed among students who initially had lower results, while among students who immediately showed high results, the growth did not exceed the margin of error.
Figure 7. Level of logical thinking. The horizontal axis shows the numbers of the weeks when the cases were solved. The vertical axis shows scores received by students.

### Table 5. Calculations for paired samples by t-test.

| Task type | Average value difference | Standard deviation of differences | Number of respondents | Empirical value of the t-test | Significance level |
|-----------|--------------------------|----------------------------------|-----------------------|-----------------------------|------------------|
| 1         | 0.79                     | 0.81                             | 32                    | 5.517                       | 0.001            |
| 2         | 0.438                    | 1.294                            | 32                    | 1.913                       | 0.1              |
| 3         | 0.672                    | 1.856                            | 32                    | 2.048                       | 0.05             |

Working with natural-scientific cases revealed that students experience perceptional difficulties dealing with tasks that do not have an unambiguous answer (for example, evaluating how the principles of naming chemical elements have changed over time using the PT, which indicates the sources of the names of elements) and that involve reasoning (for example, establishing the maximum number of connections between the synthesis of helium inside the Sun, photosynthesis and the human body). This might be probably connected to the fact that when studying natural sciences students are accustomed to looking for the correct solution, which do not require argumentation in the presence of only necessary and sufficient reference data.

The assessment study was conducted on a small sample during one semester and cannot serve as direct evidence of the efficiency of the case method for the development of logical thinking in the study of natural sciences. However, it clearly shows the possibilities of using natural science cases for routine monitoring and timely correction of the teacher’s strategy in respect to student’s logical skills. The use of the proposed tool transforms formation of those skills from spontaneous into the planned and regulated one, as well as makes it possible to objectively compare various methods of thinking development.

### 5. Conclusion

One of the important tasks of integrated natural science courses is the development of students’ logical thinking. The lack of suitable measuring instruments hinders the performance of this task, preventing teachers from systematically tracking the results of their innovations in this area and making timely changes to the methodology of teaching.

This study offers a new tool for tracking logical thinking skills for science educators and researchers. It consists of tasks of natural science cases of three types: on the continuation of the chain of reasoning, on the argumentation and on the construction of an algorithm. The advantages of this tool include the fact that, unlike the TOLT and GALT tests, it is not limited to choosing an answer from several possible ones, and, unlike a set of logical problems, it is not a “foreign” for natural science courses.

The sample of students from two universities (MPGU and RUDN University) who took part in the experiment from 2014 to 2019 showed that natural-scientific cases have a high correlation coefficient with existing tools for assessing logical thinking (GALT - 0.61 TOLT - 0.82, set of logical problems - 0.73) and to a high degree coincide with the expert assessment of logical thinking of students (correlation coefficient equals 0.87).

Natural-scientific cases fit organically into the structure of such integrated disciplines as “Natural scientific picture of the world”, “Concepts of modern natural science”, “Natural scientific project”, “Methodology of scientific cognition”. Using the example of logical thinking, the study shows that case problems can serve not only the purpose of students’ development, but also - with a certain selection of
tasks and methods of their processing - the purpose of diagnosing current changes in the students’ critical thinking level.

References
[1] Szenes E, Tilakaratna N and Maton K 2015 The Knowledge Practices of Critical Thinking *The Palgrave Handbook of Critical Thinking in Higher Education* (New York: Palgrave Macmillan US) pp 573-91
[2] Pesic J 2007 Logical and epistemological approach to critical thinking *Psihologija* **40** 173-90
[3] Mill J S 2011 *A System of Logic, Ratiocinative and Inductive* vol 1 (Cambridge: Cambridge University Press)
[4] Bird L 2010 Logical Reasoning Ability and Student Performance in General Chemistry *J. Chem. Educ.* **87** 541-6
[5] Muzaky A F, Sunarno W and Harjana 2020 Evaluating students logical thinking ability: TPACK model as a physics learning strategy to improve students logical thinking ability *J. Phys. Conf. Ser.* **1511** 012027
[6] Bowell T and Kemp G 2005 *Critical Thinking: A Concise Guide* (Routledge)
[7] Gabbay D M, Johnson R H, Ohlbach H J and Woods J 2002 *Handbook of the Logic of Argument and Inference: The Turn Towards the Practical* (North Holland: Elsevier)
[8] Solodikhina M V 2020 *Collection of Case Problems on Integrated Natural Science Courses: Textbook* (Moscow: Prometeus)
[9] Solodikhina M V and Solodikhina A A 2019 Development of critical thinking of master’s degree students using STEM cases *Edu. Sci. J.* **21** 125-53
[10] Solodikhina M V 2017 Practical lesson on natural sciences “What human consists of?” *Physics in school* **197** 201
[11] Biggs J B and Collis K F 2014 *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)* (Academic Press)
[12] Veksler V A and Reidel’ L B 2015 Particularities of the determination of pedagogical tests validity *Novainfo.ru* **36** 207-12 https://novainfo.ru/article/3845
[13] Roadrangka V, Yeany R H and Padilla M J 1983 The construction of a group assessment of logical thinking (GALT) 56th annual meeting of the National Association for Research in Science Teaching. Dallas, Texas, April pp 5-8
[14] Tobin K G and Capie W 1981 The Development and Validation of a Group Test of Logical Thinking *Educ. Psychol. Meas.* **41** 413-23
[15] Kuzina S 2014 *Boost Your Brain Using the Sherlock Holmes Method* (Moscow: AST)
[16] Kek M Y C A and Huijser H 2011 The power of problem-based learning in developing critical thinking skills: preparing students for tomorrow’s digital futures in today’s classrooms *High. Educ. Res. Dev.* **30** 329-41