Horvat, Saša; Rodic, Dušica; Jovic, Nevena; Roncevic, Tamara; Babic-Kekez, Snežana

Validation of the strategy for determining the numerical rating of the cognitive complexity of exam items in the field of chemical kinetics

CEPS Journal 13 (2023) 4, S. 111-133

Quellenangabe/ Reference:
Horvat, Saša; Rodic, Dušica; Jovic, Nevena; Roncevic, Tamara; Babic-Kekez, Snežana: Validation of the strategy for determining the numerical rating of the cognitive complexity of exam items in the field of chemical kinetics - In: CEPS Journal 13 (2023) 4, S. 111-133 - URN: urn:nbn:de:0111-pedocs-290381 - DOI: 10.25656/01:29038; 10.26529/cepsj.1235

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-290381
https://doi.org/10.25656/01:29038

DOI: https://doi.org/10.26529/cepsj.1235

Validation of the Strategy for Determining the Numerical Rating of the Cognitive Complexity of Exam Items in the Field of Chemical Kinetics

Saša Horvat*[1], Dušica Rodić[2], Nevena Jović[3], Tamara Rončević[2] and Snežana Babić-Kekez[2]

The main goal of this study was to validate the strategy for the assessment of the cognitive complexity of chemical kinetics exam items. The strategy included three steps: 1) assessment of the difficulty of concepts, 2) assessment of distractor value. and 3) assessment of concepts' interactivity. One of the tasks was to determine whether there were misconceptions by students that might have influenced their achievement. Eighty-seven students in the first year of secondary school participated in the study. A knowledge test was used as a research instrument to assess the performance, and a five-point Likert-type scale was used to evaluate the perceived mental effort. The strategy was validated using regression analysis from which significant correlation coefficients were obtained between selected variables: students' achievement and invested mental effort (dependent variables) and a numerical rating of cognitive complexity (independent variable).

Keywords: mental effort, performance, chemical equilibrium

1 *Corresponding Author. Faculty of Sciences at University of Novi Sad, Serbia; sasa.horvat@dh.uns.ac.rs.
2 Faculty of Sciences at University of Novi Sad, Serbia.
3 Master student of chemistry education at Faculty of Sciences at University of Novi Sad, Serbia.

Potrjevanje strategije za določanje številčne ocene kognitivne zahtevnosti izpitnih nalog s področja kemijske kinetike

Saša Horvat, Dušica Rodic, Nevena Jović, Tamara Rončević in Snežana Babić-Kekez

Glavni cilj te raziskave je bil potrditev strategije za ocenjevanje kognitivne zahtevnosti izpitnih nalog iz kemijske kinetike. Postopek sestoji iz treh korakov, in sicer: 1) ocene težavnosti pojmov, 2) ocene vrednosti motečih dejavnikov, 3) ocene interaktivnosti konceptov. Ena izmed nalog je bila ugotoviti, ali so se pri učencih pojavile napačne predstave, ki bi lahko vplivale na njihov uspeh. V raziskavi je sodelovalo 87 srednješolcev 1. letnika. Preizkus znanja je bil uporabljen kot instrument zbiranja podatkov glede uspešnosti, obenem pa je bila uporabljena tudi 5-stopenjska lestvica Likertovega tipa, s katero so srednješolci ocenili zaznan miselni napor. Veljavnost je bila potrjena z regresijsko analizo; ta je pokazala statično značilne vrednosti korelacijskih koeficientov med izbranimi spremenljivkami: dosežki študentov in vložen miselni napor (odvisni spremenljivki) ter numerična ocena kognitivne zahtevnosti (neodvisna spremenljivka).

Ključne besede: miselni napor, dosežek, kemično ravnovesje

Introduction

As an experimental science, chemistry also includes computation in numerical problems and reading texts. Many students have difficulty with learning chemistry because they are required to have literacy competencies and the ability to solve numerical problems (Yunus & Ali, 2012). Teaching students to solve problems and finding an effective problem-solving strategy is an important task of education and modern research (Barczi, 2013). Students often have difficulties with problem-solving in chemistry due to insufficient knowledge, as well as the existence of alternative concepts and misconceptions (Taber, 2002). Many of the alternative concepts are related to the understanding of chemical substances, chemical kinetics, and chemical changes.

The chemical kinetics concept is included in the courses of general and physical chemistry and is closely related to thermodynamics and chemical equilibrium (Justi, 2002). Numerous misconceptions and misunderstandings of the concepts of chemical kinetics have been observed in students; many misconceptions are described in the paper by Bain & Towns (2016). Misconceptions about the law of mass action were observed in students (the rate of the chemical reaction is equal to the product of the concentration of reactants) (Çam et al., 2015). Students have problems with establishing the relationship between chemical kinetics and equilibrium concepts. Specifically, they believe that the equilibrium constant refers to the reaction rate (Sözbilir et al., 2010; Turányi & Tóth, 2013).

One reason for students' misunderstandings is that at the beginning of the study of chemistry in primary school, chemical reactions are observed through changes (colour, temperature, the appearance of bubbles, the sound of cracking, etc.). In this way, students adopt the notion that a chemical reaction has occurred. However, they are introduced to chemical equilibrium when they enter secondary school. Then a new concept appears: reversible reaction and the fact that the reaction does not have to take place to the end but can take place in the opposite direction. In Van Driel's research (2002), students see direct and reversible reactions as separate and independent. The back arrow used in illustrations of reversible reactions contributes to students viewing this reaction as two reactions. These results were confirmed with different ages of students (Banks, 1997; Gorodetsky & Gussarsky, 1986; Johnstone et al., 1977; Maskill & Cachapuz, 1989).

As a result of his research, Banerjee (1991) concluded that students directly relate the value of the equilibrium constant to the rate of a chemical reaction, believing that a higher value of the equilibrium constant means a faster direct reaction. Many students do not understand that the equilibrium constant is a thermodynamic constant and depends on temperature (Gorodetsky & Gussarsky, 1986; Hackling & Garnett, 1985). Students also believe that Le Chatelier's principle can be used to predict the value of the equilibrium constant (Banerjee, 1991). Many students misinterpret Le Chatelier's principle in the field of chemical kinetics, especially regarding the effect of temperature on the rate of a chemical reaction (Turányi & Tóth, 2013).

The inability of students to calculate the chemical equilibrium constant may be an accidental error or misconception (Kousathana & Tsaparlis, 2002). Also, what confuses students are typical examples of chemical reaction equations (Tóth, 1999). In most Serbian textbooks, reactions that are given to students to solve the problem of chemical equilibrium are:

$$H_2 + I_2 \leftrightarrows 2HI \text{ or}$$
$$aA + bB \leftrightarrows cC + dD.$$

Cognitive complexity is one of the key factors influencing the ability to identify important connections between items within a complex problem and its solution. The concept of cognitive complexity is derived from George Kelly's Theory of Personal Constructs (1955). Cognitive complexity, as a dimension of personality, was introduced by Bieri (1955). The cognitive complexity of a problem task is a complex construct comprising the objective complexity and difficulty of the task (Kalyuga, 2008; van Gog et al., 2011). When solving problem tasks with a high level of cognitive complexity, differences in achievement and the assessment of mental effort between more and less successful students could be observed (Kim et al., 2014).

Research has shown that the level of cognitive complexity can predict the achievement of a problem task (Embretson & Daniel, 2008). Based on these results, manipulations in the level of cognitive complexity of the problem task can determine the difficulty of the problem (Daniel & Embretson, 2010). As far as chemical education is concerned, cognitive complexity is positively correlated with the difficulty of the problem; at the same time, increasing the cognitive complexity of the problem increases mental effort (Horvat et al., 2016; 2017, 2020, 2021; Knaus et al., 2011; Raker et al., 2013).

The design and use of a reliable instrument to assess the level of cognitive complexity for chemical tasks are very important. The application of an instrument that provides an easy way to assess the numerical rating of cognitive complexity is essential for the quantification of cognitive requirements in solving chemistry exam items (Knaus et al., 2011). The first rubric was created by Knaus et al. (2011). The developed rubrics are based on complexity theory

(Goldreich, 2008; Pippenger, 1978), which can explain a multi-connected system, and the intrinsic cognitive load construct (Sweller et al., 2011), which can explain interactivity between them.

To calculate the numerical rating of the cognitive complexity of the task, experts estimate the number of concepts included in exam items as easy, medium, or difficult from the perspective of students. Since all the concepts that are represented in the problem task are considered, the rating of the difficulty of the task is determined by applying the rubric. After determining the difficulty rating of the concepts, the interactivity of the concepts is assessed. Interactivity is assessed when there is an interdependence of concepts within the task. It is usually assessed by calculating the number of concepts included in exam items. By adding the values of the concept difficulty rating and interactivity, an overall rating of the cognitive complexity is obtained.

As Knaus et al. (2011) pointed out, good effects associated with the use of the instruments include improved knowledge about the cognitive complexity of chemical tasks and as a means of characterisation of test content for the measurement of cognitive development. The created rubrics represent a good way to determine the cognitive complexity of exam items (Knaus et al., 2011; Raker et al., 2013).

A significant contribution of the cognitive complexity rating rubrics is reflected in the development of a methodology for the logical and structured organisation of chemical concepts, which leads to more successful learning (Segedinac et al., 2018). The construction of an invalid instrument for measuring cognitive complexity could lead to invalid test results. Therefore, the development of a valid instrument is of crucial importance.

Research problem and aim

The previously described rubrics have proven to be reliable and valid tools for application in chemistry education (Knaus et al., 2011; Raker et al., 2013). However, due to the specificity of domains in chemistry, these rubrics should be further improved. They are developed in the form of a table for assessing the difficulty of the concepts and are of great importance because they are objective and precisely defined by experts (Horvat et al., 2016; 2017, 2020, 2021). In this study, we have developed a table for the domain of Chemical Kinetics.

The main objective of this research is to validate the table for assessing cognitive complexity in chemical kinetics problems. In chemical equilibrium, only the kinetic aspect and Le Chatelier's principle were observed. To validate the proposed strategy, it was necessary to statistically confirm the dependences of students' achievement and mental effort from calculated cognitive complexity.

From the research problem and aim, the following research questions were derived:

- How does an increase in the numerical rating of cognitive complexity affect students' achievements?
- How does an increase in the numerical rating of cognitive complexity affects students' invested mental effort?
- How does an increase in students' invested mental effort affects students' achievements?
- What factors affect students' achievement?

From the defined research questions, three hypotheses were set:

- An increase in the numerical rating of cognitive complexity leads to a decrease in students' achievements and an increase in invested mental effort;
- An increase in students' invested mental effort leads to a decrease in students' achievements.;
- Students possess misconceptions that affect the achievement of the students.

Method

Participants

The field of chemical kinetics and chemical equilibrium is studied in the subject of chemistry in the first year in the textbooks of the secondary schools with general, natural-mathematical, and socio-linguistic emphases. Eighty-seven participants students in the first year of the secondary school Gymnasium 'Svetozar Marković' participated in this study. They studied the concepts of chemical kinetics during the second semester of the 2018/19 school year. The research was conducted in June, at the end of the school year. The students voluntarily took part in this study.

Instrument

As a research instrument, the Knowledge Test was used, which was specially created for the needs of the research. The time provided for the test solving was 45 minutes. The test consisted of five problem tasks. A completely solved task was scored with one point, and tasks that were partially solved were scored by item requirements. The third task contained only one item, and the fifth task had two items, while the first, second and fourth had three items each. The maximum score on the test was five points. The tasks of this test were taken from 'Zbirka zadataka iz hemije za I II razred gimnazije i srednje škole - Collection of tasks in chemistry for the first and second grade of gymnasiums and high schools (English)' (Nikolajević & Šurjanović, 2011). Also, when the procedure for solving the arithmetic task had been completely correct, and the student had made a mistake in rounding the decimal numbers, such a task was scored correctly because it was considered that the student had completely mastered the concepts represented in this task.

The knowledge test also served as an instrument for the evaluation of invested mental effort by including a 5-point Likert scale, as proposed by Knaus et al. (2011) and Raker et al. (2013). During the statistical processing of the results, the appropriate numerical values were assigned for descriptive estimates of the assessment. Specifically, 'very easy', 'easy', 'neither difficult nor easy', 'difficult', and 'very difficult' were represented as numerical values 1, 2, 3, 4, and 5, respectively.

The obtained results were processed by the statistical software program IBM SPSS Statistics 24.

Instrument validation

Test quality was assessed by pre-test and post-test quality assurance parameters. The model was described by Segedinac et al. (2011). Pre-test quality assurance parameters had been determined by four experts whose specific field is chemistry education. The test was assessed as valid as the tasks were concurrent with the subject syllabus and recommended textbooks. The test tasks were assessed by the experts as diverse, with precisely clearly established requirements and meaningful sentences that satisfy the linguistic standards.

Basic statistical parameters: reliability coefficient, task discrimination index, test discrimination index, task difficulty index and test difficulty index were used as post-test assurance parameters. Reliability was calculated as a measure of internal consistency and was expressed as the Cronbach α coefficient. A value of Cronbach α coefficient of 0.58 for achievement represents a satisfactory coefficient of reliability (Taber, 2018) as it was appropriate for the present number of exam items (Loewenthal, 2004; Moss et al., 1998; Tavakol & Dennick, 2011; Taber, 2018). For self-assessed mental effort, the value of the Cronbach α coefficient was 0.71, which indicated high reliability (Jonsson & Svingby, 2007; Taber, 2018). The task difficulty index was calculated as the average achievement on the task, while the test discrimination index was calculated as the average achievement of all six tasks on the test. A rule of thumb was used to interpret the values (Towns, 2014). Task difficulty indices were ranging from 6.90% to 34.48% (the average value is 22.53%, and the test was characterised as difficult). Two tasks had a difficulty index of 25–75%, which makes them moderate tasks (Pande et al., 2013; Towns, 2014). Three tasks had a difficulty index of less than 25% and were categorised as difficult tasks. When observed, the difficulty of items was in the range of 0% to 54.02%. Eight items had a difficulty index of less than 25%, while four items had difficulty indices of 25–75%. Task and item discrimination indices were calculated using the extreme group method. The sample was divided into two groups using the average score, creating an upper half and a lower half. Item discrimination indices were obtained from the subtraction of the average score of 30% of students in the upper half and lower half (Towns, 2014). The test discrimination index was obtained as the average value of all single-task discrimination indices. Task discrimination indices had values from 0.17 to 0.70 (mean 0.50 describes an excellent discrimination index). Four tasks were characterised as excellent based on a discrimination index higher than 0.4. One task had a poor discrimination index of 0.17 (Towns, 2014; Zubairi & Kassim, 2006). This task should be revised or discarded for the next use. The discrimination indices of items were in the range of 0 to 0.91. If we consider parts of tasks (i.e., items), seven of them have an excellent discrimination index (higher than 0.4), four of them had a discrimination index in the range of 0.2 to 0.4, which were good items that can be improved as needed, while only one item had a poor discrimination index of 0 and should be rejected. The created test used in this research, as far as the post-test quality guarantee assurance parameters were concerned, showed good metric characteristics.

The mean value of the mental effort is 3.21, which means that the test can be characterised as 'neither difficult nor easy' according to the applied Likert scale. The basic statistical parameters of the test for student achievement are presented in Table 1.

Table 1

Basic statistical parameters

Parameter	Students' achievement[1]	Students' ratings of mental effort[2]
Average	1.53	3.21
Standard deviation	1.07	0.69
Standard skewness	0.82	-0.45
Standard kurtosis	0.21	2.87
Minimum	0.00	1.00
Maximum	4.17	5.00
Range	4.17	4.00

Note. [1]Achievement range 0–5. [2]Mental effort range 1–5.

The kurtosis and skewness values were considered, indicating a normal distribution of achievements at the level of trust of 95%. However, this was not confirmed with an additional Kolmogorov-Smirnov test (D = 0.14; p < .001). The analysis of z-scores or nontypical values (outliers) revealed the existence of several scores greater than ± 2.58 of the maximum 1% of allowed z-scores, which additionally cannot satisfy the criterion of normal distribution (Mayers, 2013). The values of the coefficients of skewness and kurtosis show that the normal distribution is not present in the mental effort estimates. The Kolmogorov-Smirnov test (D = 0.18; p < .001) also confirmed this assumption that the normal distribution was not presented.

The test was validated by observing the relation between students' achievement and invested mental effort. As neither the achievement nor the mental effort of the respondents satisfies the normal distribution, the dependence of the student's achievement from the average mental effort through the non-parametric Spearman's ρ-coefficient was observed.

This dependence is described by the equation Achievement= 0.37-0.04×Mental effort; it is described as a moderate correlation (r_s = -0.29; p = 0.01). The higher invested mental effort cause a decrease in students' achievements. Cohen's d-effect size is 'much larger than typical' (1,86), so an attempt to replicate the study with a larger sample may be justified (Morgan et al., 2011). This had already been confirmed in studies that had the topic of validation of procedures for the assessment of problems' cognitive complexity (Horvat et al., 2016, 2017, 2020, 2021; Knaus et al., 2011; Raker et al., 2013).

Research design

Based on the set research aim and problems, the research tasks were defined as follows:

- Creation of a table for assessing the difficulty of the concepts of exam items in the field of chemical kinetics and chemical equilibrium;
- Determining the numerical rating of the difficulty of the concepts included in the exam items of the created test;
- Conducting a test;
- Statistical processing of results; and
- Analysis of test results.

Additionally, after a detailed analysis of the students' answers on test items, some misconceptions that have had influenced students' test results were identified and analysed.

Results

Table creation

To fulfil the aim of this research, a table for assessing the difficulty of the concepts represented in the area of chemical kinetics and chemical equilibrium has been created. In addition to the concepts that can be represented in the problems of chemical kinetics and chemical equilibrium and their difficulty level, this table contains distractors and an assessment of the interactivity of concepts. The created table is represented in Table 2. Chemical equilibrium was observed only from the kinetic side and not from the thermodynamic side. The principle for using this table is simple and objective. The total cognitive complexity of the task is expressed as a numerical value.

Table 2

Table for assessing the difficulty of the concepts in chemical kinetics problems

CHEMICAL KINETICS	Difficulty
An expression for the rate of a chemical reaction and the law of mass action	easy
Calculation of the rate of a chemical reaction	medium
Calculation of the rate of a chemical reaction by applying the law of mass action	difficult
CHEMICAL EQUILIBRIUM	
Expression for the equilibrium constant, Le Chatelier's principle	easy
Calculation of equilibrium constant and equilibrium concentrations	medium
Calculation of the initial concentrations of reactants	difficult

ADDITIONAL CONCEPTS	
Homogeneous reactions - class of chemical reactions that occur in a single gaseous phase (g) or aqueous phases (aq)	0
Heterogeneous reactions - class of chemical reactions where one of the reactants occurs in a liquid (l) or a solid (s)	1
INTERACTIVITY	
Task consists of 2 concepts	0
Task consists 3 concepts	1
Task consists of more than 3 concepts	2

All the concepts needed to solve problem tasks in the field of chemical kinetics are assessed as easy, medium-difficult, or difficult. The difficulty ratings of concepts were determined through a panel discussion by a group of four experts, all of whom were university professors. The assessment of the difficulty of each concept was assessed from the student's perspective. The experts assessed the difficulty of each concept independently. All disagreements in the assessment between experts were eliminated through discussion.

The numerical difficulty rating was obtained by estimating the difficulty of each represented concept and using the rubrics by Knaus et al. (2011) and Raker et al. (2013). The concept of chemical kinetics can be assessed as easy, medium, or difficult. The concept was considered 'easy' if the student needed to write an expression for the rate of a chemical reaction for a particular reactant or reaction product. Also, writing an expression for the rate of a chemical reaction by applying the law of mass action was an easy concept. If the task aim was to calculate the rate of the chemical reaction by varying the concentrations of the reactants and products of the reaction over time, the concept was considered as a 'medium' difficulty concept. A 'difficult' concept implied that the student could successfully apply the law of mass action and calculate the chemical reaction rate constant to determine the change in the rate of a chemical reaction depending on the change in concentration or change in pressure in the chemical reaction.

Another group of concepts represented in Table 3 are concepts of chemical equilibrium. This concept was considered 'easy' if students were asked to write an expression for the equilibrium constant of a chemical reaction. The application of Le Chatelier's principle (i.e., the influence of change on the chemical equilibrium: concentration, pressure or temperature) was considered an 'easy' concept. If students were requested to calculate the chemical equilibrium constant or equilibrium concentration of one of the participants in the reaction, it was considered a 'medium-difficult' concept. A 'difficult' concept involves calculations with initial concentrations of reactants.

In the tasks, additional concepts were considered as the aggregate state of the participants in the chemical reaction. Reactants that were in the gaseous state (g) or as solutes (aq) are included in the expressions for calculating the equilibrium constant and the expressions for the rate of the chemical reaction as a rule. Reactants in the liquid (l) or solid-state (s) were not taken into account. If the participants in the chemical reaction were in a gaseous state or were dissolved substances (homogeneous reaction), the value of additional concepts is 0. However, if one of the substances was solid or in a liquid state (heterogeneous reaction), its concentration or partial pressure could not be changed; the additional concept had a value of 1. In this manner, a student's lucky guess was eliminated, because if he omitted a reactant that is in a liquid or solid state, it could be considered that he has mastered the concepts of chemical kinetics and equilibrium.

Adding to the difficulty of the concepts, the value of distractors' final step for the assessment of cognitive complexity was the determination of interactivity. It was assessed based on the number of concepts. In problems with two concepts, the value of interactivity was 0; in problems with three concepts, it had 1; it had a value of 2, in problems with more than three concepts.

Cognitive complexity determination

After estimating the difficulty of the concepts represented in the exam items, determining the value of additional concepts and evaluating the interactivity, a numerical rating of cognitive complexity was obtained using the method proposed by Knaus et al. (2011) and Raker et al. (2013). A concrete example is shown below:

In the reaction system $CO(g) + Cl_2(g) \rightarrow COCl_2(g)$, the CO concentration was increased from .3 to 1.2 mol/dm^3 and the chlorine concentration from .2 to .6 mol/dm^3. How many times did the reaction rate increase?

This task contains only one item, and it can be solved in three simple steps: calculating the rate of the chemical reaction at the beginning of the chemical reaction, then calculating the rate of the chemical reaction after a change in concentrations, and finally determining the ratio of the rates of chemical reactions. This task contains two concepts from chemical kinetics. A numerical rating of cognitive complexity was obtained in this way:

- Law of mass action - a concept that is 'easy' according to Table 3 and had the value of the difficulty of concept 1 according to the method by Knaus et al. (2011) and Raker et al. (2013);
- Calculation of the rate of a chemical reaction by applying the law of mass action - difficulty concept according to Table 3 and has the value of the

difficulty of concept 4 according to the method by Knaus et al. (2011) and Raker et al. (2013);

- Since it is a homogeneous reaction, the additional concept value has a value of 0;
- The problem has two concepts, so the value of interactivity is 0.

By adding all the numerical values, the overall rating of cognitive complexity of 5 was obtained.

The following example is a task in the field of chemical equilibrium:

In system A(aq) + 2B (aq) ⇆ C(aq) + D(l) equilibrium concentrations are [A]=0.06 mol/dm³, [B]=1.2 mol/dm³, [C]=2.16 mol/dm³, [D]=2.16 mol/dm³. Calculate the equilibrium constant of this reaction as well as the initial concentrations of the reactants.

This task has two items. The first item involves the calculation of the equilibrium constant. When calculating the equilibrium constant, students need to pay attention because the expression does not take into account the concentration of product D, which is in a liquid state. At the beginning of the task, it can be seen that a heterogeneous reaction is present in this task. The second item involves the calculation of starting concentrations. The initial concentrations of the products are always zero because they have not formed yet in a chemical reaction. The initial concentrations of the reactants are obtained by adding the equilibrium concentrations to the participating concentration in the reaction.

Based on Table 3, in the same manner as in the previous example, the difficulty was determined. The concepts represented in the task are as follows:

- Calculation of equilibrium constant - a concept that is of medium difficulty according to Table 3 and has the value of the difficulty of concept 2 according to the method of Raker et al. (2013); and
- Calculation of the initial concentrations of reactants - a concept considered difficult according to Table 3 and has the value of the difficulty of concept 4 according to the method by Knaus et al. (2011) and Raker et al. (2013).
- This task has an additional concept that has a value of 1 and two concepts are represented, so the interactivity was evaluated with 0.

By adding all the numerical values, a numerical rating of cognitive complexity of 7 was obtained. Likewise, the numerical rating of cognitive complexity was successfully assessed for other exam items represented in the test. The calculated values were in the range of 4 to 7.

Procedure validation

As the distributions of students' achievements and mental effort did not satisfy the criteria of normal distribution, validation of this procedure is not possible by linear regression. Therefore, instead of the Pearson coefficient, the correlation between the variables was performed using the Spearman ρ-correlation coefficient, which is used when the distribution is not normal. All 435 items were observed. The number of items was obtained by multiplying the number of students by the number of tasks. In the first phase of validation, the dependence of students' achievement (independent variable) on cognitive complexity (dependent variable) was observed.

The obtained coefficients ($r_s = -0.33$; $p < .001$) indicate a moderate correlation between examined variables (Evans, 1996). The correlation of this dependence is Achievement = .70 - .09 × Cognitive complexity. Although the correlation coefficient is small and Cohen's d-effect size value is larger than typical (3.28), the correlation is significant because it was made on the basis of the correlation of a large number of variables (Brace et al., 2006; Cohen, 1988; Morgan et al., 2011), so it is statistically significant considering the number of items.

The next step of the procedure validation was to examine whether there is a correlation between students' self-invested mental effort and cognitive complexity. The correlation of this dependence is Mental effort = 2.10 + 0.21 × Cognitive complexity.

The obtained coefficients ($r_s = .22$; $p < .001$) indicate a weak but statistically significant correlation between examined variables (Evans, 1996). Cohen's d effect size value shows a larger effect size than typical (1.75), and it is a significant correlation because it was made on the basis of the correlation of a large number of variables (Morgan et al., 2011).

Misconception identification

The cause of the low achievement values can also be misconceptions that were observed in students who participated in the research. Tests have been used for many years as assessment tools for the identification of students' misconceptions in science. The test questions developed for this purpose thus far are available in many forms, such as interview, multiple-choice questions, open-ended questions, multi-tier questions, and others (Soeharto et al., 2019). Tests such as multiple-choice tests (32.23%) and multiple-tier tests (33.06%) are used as diagnostic tools for the identification of misconceptions in more than 65% of research papers. In contrast, open-ended questions give students the freedom to think and write their ideas but also cause difficulties in interpreting and analysing student answers, as some response answers may not be useful, and in reviewing answers, which might be time-consuming (Soeharto et al., 2019, Soeharto & Csapó, 2021). When the frequency of wrong answers is higher than 10%, it can be considered a misconception (Yan & Subramaniam, 2018).

The first task contained three items and required writing expressions for equilibrium constants for given chemical reactions with stoichiometric coefficients. The achievement by items was 24% in the first and third items and 54% in the second. The reason for this difference in achievement is precisely the misconception observed in students. In 21.84% of the respondents, it was observed that they did not include participants that are chemical elements in the expression for the equilibrium constant. This is in line with the previous research (Çakmakci, 2010) in which it was observed that students mix the concepts of enthalpy and the rate of a chemical reaction. Specifically, the students are probably confused by the fact that an element in its standard state has a standard enthalpy of formation of 0 kJ/mol, and they are not entered into the expressions for the calculation of the enthalpy of a chemical reaction. It was also seen that when students were writing the expression for the equilibrium constant, the equilibrium concentration did not agree with the stoichiometric coefficient (18.39%), and instead of multiplying, the students added the equilibrium concentrations of reactants and reaction products (12.64% of students). This is also related to the mixing of the concept of enthalpy and chemical kinetics. I.e., in the curriculum, chemical kinetics comes immediately after thermochemistry teaching units, so probably many of the students fail to observe these two concepts separately.

The second task also contained three items. This task required writing expressions for the rate of direct chemical reactions by applying the law of mass action. The achievement by items was 21% in the first, 6% in the second and 15% in the third item. This task contained additional concepts in all three items, specifically that chemical reactions were heterogeneous reactions. The low achievement is caused precisely by the fact that the students included in the expressions for the rate of the chemical reaction the participants of the chemical reaction that are in a liquid or solid-state (Kousathana & Tsaparlis, 2002). This misconception was observed in 14.94% of students. In this task, instead of writing expressions for the rate of a chemical reaction, students wrote expressions for the equilibrium constant (19.54% of students) or an expression for the rate of a chemical reaction by changing the concentration of reactants per unit time (16.09% of students). This result had previously been mentioned in numerous studies in which it was seen that students combine the concepts of chemical kinetics and equilibrium, believe that the equilibrium constant refers to the rate of a chemical reaction, and do not understand the law of mass action (Çam et al., 2015; Sözbilir et al., 2010; Turányi & Tóth, 2013 cited in Bain & Towns, 2016). The lowest achievement was in the second item, for which students were required to write the rate of a chemical reaction by applying the law of mass action to the following chemical reaction:

$$2\ H_2O_2(l) \leftrightarrows 2\ H_2O(l) + O_2(g)$$

Since hydrogen peroxide is a liquid substance, according to the law of mass action, its concentration cannot be changed, so the rate of a direct chemical reaction is equal to the rate constant of the chemical reaction:

$$v = k.$$

However, in 21.83% of students, a misconception was observed that they expressed the rate of direct chemical reaction as the rate of reverse chemical reaction:

$$v = k[O_2]$$

where it can be seen that the students calculated with the oxygen concentration probably because it was the only participant in the reaction in the gaseous state. Students' problems with applying the law of mass action in chemical kinetics have long been observed (BouJaoude,1993). This misconception may be because students think that the rate of direct and reverse chemical reactions is equal (Bain & Towns, 2016; Cliff, 2009; Hackling & Garnett, 1985;).

The third task had one item and the average achievement was 24%. This was a computational task in which the students needed to calculate how many times the rate of a chemical reaction had changed by applying the law of mass action.

The fourth task contained three items, requiring students to determine how an increase in temperature (item 1), a decrease in pressure (item 2), and an increase in the concentration of one of the reactants (item 3) affect the concentration of the reaction product of the exothermic chemical reaction. The achievement by items was 29%, 39%, and 36%, respectively. The misconception that an increase in temperature favours an exothermic chemical reaction had already been observed many times in previous research, which can be found in Bain and Towns (2016). This misconception was observed in 24.13% of students who participated in the study.

The fifth task contained two items. The first item referred to the calculation of the value of the equilibrium constant in a chemical reaction when one reaction product is in a liquid state. The achievement on this item was 14%. The second item concerned the calculation of the initial concentrations of the reactants. The achievement on this item was 0%. This is cognitively the most complex task, the task with the lowest achievement of the respondents and the highest amount of mental effort of the students. Many of the students did not even attempt to do this task. It has been noted that students' problems also occur when calculating initial concentrations (Kalainoff et al., 2012).

If we recall the research questions and hypotheses, we can conclude that they were justified. Increasing the numerical rating of cognitive complexity has led to lower student achievement and has imposed higher mental effort on students, which is fully consistent with the results of the previously published research (Horvat et al., 2016, 2017, 2020, 2021; Knaus et al., 2011; Raker et al., 2013). Careful selection of tasks, with the gradual introduction of new concepts that students have just mastered, can lead to a more permanent formation of knowledge structures and deeper understanding. Using this procedure, teachers can gradually make problems more complex. Thus, they can foster the development of students' problem-solving skills while preventing the efficiency of the teaching process from decreasing at any time due to the overload of working memory. Designing teaching materials of different levels of complexity is a better way to assess learning outcomes and re-examine cognitive load through mental effort measures. It is also necessary for teachers to consider problem-solving strategies used by their students because the numerical solution of the problem does not provide insight into the knowledge of concepts and students' understanding of concepts (BouJaoude & Barakat, 2000). This procedure and the test provide us with valuable information: by carefully reviewing students' answers to solved tasks, teachers can see which concepts represent problems for students and modify the already created table and procedure to attempt to enhance students' achievements.

Conclusions and implications

The aim of this research was first to create and then to validate a strategy for the assessment of the numerical rating of cognitive complexity in the domain of chemical kinetics. The first phase of validation was the validation of the instrument itself, which showed a good metric characteristic in terms of the discrimination index (0.50). Four tasks had an excellent discrimination index, while one task had a poor discrimination index. The test was difficult, and the test difficulty index was 22.53%.

The second phase of validation was to examine the correlations between students' achievement and their self-perceived mental effort from the numerical rating of the cognitive complexity of the exam items. The cognitive complexity of the problem was assessed by creating a table for assessing the difficulty of the concepts in chemical kinetics problems. In addition to difficulty and interactivity, this table contained additional concepts; it had previously been developed, and the numerical difficulty ratings were calculated using the cognitive complexity-rating rubric, which had previously been developed (Knaus et al., 2011; Raker et al., 2013). Since students' achievement and mental effort do not satisfy the normal distribution, the validity of the procedure was performed by correlation analysis via nonparametric Spearman's ρ-coefficient, with the obtained values of the correlation coefficient indicating a weak and moderate correlation between variables that is statistically significant.

The development of such a strategy for the assessment of the numerical rating of cognitive complexity enables the assessment of the difficulty of concepts included in exam items in different chemistry domains, such as the chemical kinetic one in this research. In this way, teachers can easily estimate the cognitive complexity of exam items, which allows them to control the mastered concepts and the complexity of the problem. This procedure makes it possible to gradually adopt chemical concepts without overloading students' working memory and, at the same time, to achieve the best possible achievement. By analysing the students' achievement, it was observed that the achievement decreases with the increase of the numerical rating of cognitive complexity, and at the same time, mental effort is increased.

Numerous misconceptions on the part of students were seen in the analysis of their results. Some of them were that students mixed the concepts of enthalpy and kinetics, did not understand the law of mass action, misunderstood the concepts of exothermic reactions, and they were not skilled at solving numerical problems.

A further direction of research could be the understanding of students' misconceptions and the possibility of correcting them. This could be done by applying multi-layered tests, interviews with students or by applying an eye-tracking technique.

Acknowledgements

The authors acknowledge the financial support of the Ministry of Education, Science and Technological Development of the Republic of Serbia (Grant No. 451-03-9/2021-14/ 200125)".

References

Bain, K., & Towns, M. (2016). A review of research on the teaching and learning of chemical kinetics. *Chemistry Education Research and Practice*, *17*(2), 246–262. https://doi.org/10.1039/C5RP00176E

Banerjee, A. (1991) Misconceptions of students and teachers in chemical equilibrium. *International Journal of Science Education*, *13*(3), 355–362. https://doi.org/10.1080/0950069910130411

Banks, P. J. (1997*). Students' understanding of chemical equilibrium.* [Unpublished MA thesis]. Department of Educational Studies, University of York.

Barczi, K. (2013). Applying Cooperative techniques in teaching problem solving. *Center for Educational Policy Studies Journal, 3*(4), 61–78. https://doi.org/10.26529/cepsj.223

Bieri, J. (1955). Cognitive complexity–simplicity and predictive behaviour. *Journal of Abnormal and Social Psychology, 51*(2), 263–268.

BouJaoude, S., & Barakat, H. (2000). Secondary school students' difficulties with stoichiometry, *School Science Review*, *81*(296), 91–98.

Brace, N., Kemp, R., & Snelgar, R. (2006) *SPSS for Psychologists: A guide to data analysis using SPSS for Windows* (3rd ed.). Routledge.

Çakmakci, G. (2010). Identifying alternative conceptions of chemical kinetics among secondary school and undergraduate students in Turkey. *Journal of Chemical Education*, *87(4)*, 449–455. https://doi.org/10.1021/ed8001336

Çam, A., Topçu, M. S., & Sülün, Y. (2015). Preservice science teachers' attitudes towards chemistry and misconceptions about chemical kinetics. *Asia-Pacific Forum on Science Learning and Teaching*, *16*(2), 1–6.

Cliff, W. H. (2009). Chemistry misconceptions associated with understanding calcium and phosphate homeostasis. *Advance in Physiology Education*, *33*(4), 323–328. https://doi.org/10.1152/advan.00073.2009

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.

Daniel, R. C., & Embretson, S. E. (2010). Designing cognitive complexity in mathematical problem-solving items. *Applied Psychological Measurement, 34*(5), 348–364. https://doi.org/10.1177/0146621609349801

Embretson, S. E., & Daniel, R. C. (2008). Understanding and quantifying cognitive complexity level in mathematical problem solving items. *Psychological Science, 50*, 328–344.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Brooks/Cole.

Goldreich, O. (2008). *Computational complexity: A conceptual perspective*. Cambridge University Press.

Gorodetsky, M., & Gussarsky, E. (1986) Misconceptualisation of the chemical equilibrium concept as revealed by different evaluation methods. *European Journal of Science Education*, 8*(4),* 427–441. https://doi.org/10.1080/0140528860080409

Hackling, M. W., & Garnett, P. (1985). Misconceptions of chemical equilibria. *European Journal of Science Education*, *7*(2), 205–214. https://doi.org/10.1080/0140528850070211

Horvat, S. A, Mihajlović, J., Rončević, T. N., & Rodić, D. D. (2021). Procedure for the assessment of cognitive complexity: Development and implementation in the topic „Hydrolysis of Salts“. *Macedonian Journal of Chemistry and Chemical Engineering*, *40*(1), 119–130. https://doi.org/10.20450/mjcce.2021.2240

Horvat, S. A. , Rončević, T. N., Arsenović, D. Z , Rodić, D. D., & Segedinac, M. D. (2020). Validation of the procedure for the assessment of cognitive complexity of chemical technology problem tasks. *Journal of Baltic Science Education*, *19*(1), 64–75. https://doi.org/10.33225/jbse/20.19.64

Horvat, S., Rodić, D. D., Segedinac, M. D., & Rončević, T. N. (2017). Evaluation of cognitive complexity of tasks for the topic hydrogen exponentin the solutions of acids and bases. *Journal of Subject Didactics*, *2*(1), 33–45. https://doi.org/10.5281/zenodo.1238972

Horvat, S., Segedinac, M. D., Milenković, D. D., & Hrin, T. N. (2016). Development of procedure for the assessment of cognitive complexity of stoichiometric tasks. *Macedonian Journal of Chemistry and Chemical Engineering*, *35*(2), 275–284. https://doi.org/10.20450/mjcce.2016.893

Johnstone, A. H., Macdonald, J. J., & Webb, G. (1977). Chemical equilibria and its conceptual difficulties, *Education in Chemistry*, *14*(6), 169– 171.

Jonsson, A., & Svingby, G. (2002). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research and Reviews*, *2*(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Justi, R. (2002), Teaching and learning chemical kinetics. In J. K. Gilbert, De O. Jong, R. Justi, D. Treagust , & J. H. Van Driel (Eds.), *Chemical Education: Towards Research based Practice* (pp. 293–315). Kluwer.

Kalainoff, M., Lachance, R., Riegner, D., & Biaglow, A. A. (2012). Computer algebra approach to solving chemical equilibria in general chemistry. *PRIMUS*, *22*(4), 284–302.

Kalyuga, S. (2008). *Managing cognitive load in adaptive multimedia learning. information science reference*. Hershey.

Kelly, G. A. (1955). *The psychology of personal construct. A theory of personality*. Taylor & Francis.

Kim, S. J., Aleven, V., & Dey, A. K. (2014). Understanding expert-novice differences in geometry problem-solving tasks. *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. https://doi.org/10.1145/2559206.2581248

Knaus, K., Murphy, K., Blecking, A., & Holme, T. (2011). A valid and reliable instrument for cognitive complexity rating assignment of chemistry exam items. *Journal of Chemistry Education*, *88*(5), 554–560. https://doi.org/10.1021/ed900070y

Kousathana, M., & Tsaparlis, G., (2002). Students' errors in solving numerical chemical equilibrium problems. *Chemistry Education Research and Practice*, *3*(1), 5–17 https://doi.org/10.1039/B0RP90030C

Loewenthal, K. M. & Lewis, C. A. (2001). *An introduction to psychological tests and scales*. Psychology Press. https://doi.org/10.4324/9781315782980

Maskill, R., & Cachapuz, A. F. C. (1989). Learning about the chemistry topic of equilibrium: The use of word association tests to detect developing conceptualisations. *International Journal of Science Education*, *11*(1), 57–69. https://doi.org/10.1080/0950069890110106

Mayers, A. (2013). *Introduction to statistics and SPSS in psychology*. Pearson Education.

Morgan, G. A., Leech, N. L., Gloackener, G. W., & Barret, K. C. (2011). *BM SPSS for introductory statistics: Use and interpretation*. Routledge.

Moss, S., Prosser, H., Costello, H., Simpson, N., Patel, P., Rowe, S.,Tuner, S., & Hatton, C. (1998). Reliability and validity of the PAS-ADD checklist for detecting psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research*, *42*(2), 173–183. https://doi.org/10.1046/j.1365-2788.1998.00116.x

Pande, S. S., Pande, R. P., Parate, V. P., Nikam, A. N., & Agrekar, S. H. (2013). Correlation between difficulty and dis-crimination indices of MCQs in formative exam in physiology. *South-East Asian Journal of Medical Education, 7*(1), 45–50. https://doi.org/10.4038/seajme.v7i1.149

Pippenger, N. (1978). Complexity theory. *Scientific American, 238*(6), 114–125.

Raker, J. R., Trate, J. M., Holme, T. A.. & Murphy, K. (2013). Adaptation of an instrument for measuring the cognitive complexity of organic chemistry exam Items. *Journal of Chemical Education*, *90*(10), 1290–1295. https://doi.org/10.1021/ed400373c

Segedinac, M., Segedinac, M., Konjović, Z., & Savić, G. (2011). A formal approach to organization of educational objectives. *Psihologija*, *44*(4), 307–323. https://doi.org/10.2298/PSI1104307S

Segedinac, M. T., Horvat, S., Rodić, D. D., Rončević, T. N., & Savić, G. (2018). Using knowledge space theory to compare expected and real knowledge spaces in learning stoichiometry. *Chemistry Education Research and Practice*, *19*(3), 670–680. https://doi.org/10.1039/C8RP00052BC

Soeharto, S., Csapó, B., Sarimanah, E., Dewi, F. I., & Sabri, T. (2019). A review of students' common misconceptions in science and their diagnostic assessment tools. *Jurnal Pendidikan IPA Indonesia, 8*(2), 247–266.

Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon, 7*(11), e08352. https://doi.org/10.1016/j.heliyon.2021.e08352

Sözbilir M., Pinarbasi T., & Canpolar N. (2010). Prospective chemistry teachers' conceptions of chemical thermodynamics and kinetics. *Eurasian Journal of Mathematics, Science and Technology Education*, *6*(2), 111–120. https://doi.org/10.12973/ejmste/75232

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Šurjanović, M., & Nikolajević, R. (2011). *Hemija - Zbirka zadataka iz hemije - za 1. i 2. razred gimnazije prirodno-matematičkog smera, medicinsku i poljoprivrednu školu* [Chemistry - Collection of tasks in chemistry - for the 1st and 2nd grade of the gymnasium of natural mathematics, medical and agricultural school]. Zavod za udžbenike i nastavna sredstva.

Taber, K. S., (2002). *Chemical misconceptions – prevention, diagnosis and cure: Theoretical background.* Royal Society of Chemistry.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of*

Medical Education, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tóth, Z. (1999), Egy kémiai tévképzet nyomában. Az egyensúlyi állandó bevezetésének lehetöségei és problémái [Tracing a chemical misconception. The challenges and problems of the introduction of the chemical equilibrium constant]. *Iskolakultúra*, *9*, 108–112.

Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education, 91*(9), 1426–1431. https://doi.org/10.1021/ed500076x

Turányi, T., & Tóth, Z., (2013). Hungarian university students' misunderstandings in thermodynamics and chemical kinetics. *Chemistry Education Research and Practice*, *14*(1), 105–116. https://doi.org/10.1039/C2RP20015E

Van Driel, J. H. (2002) Students' corpuscular conceptions in the context of chemical equilibrium and chemical kinetics. *Chemistry Education: Research and Practice, 3*(2), 201–213. https://doi.org/10.1039/B2RP90016E

Van Gog, T., Kester., L., & Paas, F. (2011). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology, 25*(4), 584–587. https://doi.org/10.1002/acp.1726

Yan, Y. K., & Subramaniam, R. (2018). Using a multi-tier diagnostic test to explore the nature of students' alternative conceptions on reaction kinetics. *Chemistry Education Research and Practice*, *19*(1), 213–226 https://doi.org/10.1039/C7RP00143F

Yunus, W. M. D. Z. W., & Ali, Z. M. (2012). Urban Students' Attitude towards Learning Chemistry. *Procedia-Social and Behavioral Sciences*, *68*, 295–304. https://10.1016/j.sbspro.2012.12.228

Zubairi, A., Lide, N., & Abu Kassim, N. L. (2006). Classical And Rasch Analyses Of Dichotomously Scored Reading Comprehension Test Items. *Malaysian Journal of ELT Research, 2*, 1–20.

Biographical note

Saša Horvat, PhD, is an assistant professor in the field of chemistry education on the Faculty of Sciences at University of Novi Sad, Serbia. His research interests include problem solving in chemistry, cognitive complexity, triplet model of knowledge representation, systemic approach to teaching and learning.

Dušica Rodić, PhD, is an associate professor in the field of chemistry education on the Faculty of Sciences at University of Novi Sad, Serbia. Her research interests include chemistry triplet relationship, evaluation of mental effort, cognitive complexity, development of multi-tier tests, systemic approach in chemistry teaching, knowledge space theory applications and misconceptions in chemistry.

Nevena Jović, PhD, is master student of chemistry education on the Faculty of Sciences at University of Novi Sad, Republic of Serbia. Her research interests include problem solving and problem cognitive complexity.

Tamara Rončević, PhD, is an assistant professor in the field of chemistry education on the Faculty of Sciences at University of Novi Sad, Republic of Serbia. Her research interests include systemic approach to teaching and learning chemistry, systems thinking, illustrative methods in teaching chemistry, cognitive load theory applied within chemistry education.

Snežana Babić Kekez, PhD, is a full professor in the field of pedagogy on the Faculty of Sciences at University of Novi Sad, Republic of Serbia. Her research interests include moral education, educational policy, population education and development of pedagogical culture with parents.