



J. Serb. Chem. Soc. 81 (12) 1455–1471 (2016)
JSCS–4941

Enhancement and assessment of students' systems thinking skills by application of systemic synthesis questions in the organic chemistry course

TAMARA N. HRIN*#, DUŠICA D. MILENKOVIĆ#, MIRJANA D. SEGEDINAC#
and SAŠA HORVAT

Faculty of Sciences, University of Novi Sad, Trg D. Obradovića 3, 21000 Novi Sad, Serbia

(Received 11 August, revised 28 September, accepted 27 October 2016)

Abstract: Many studies in the field of science education have emphasized the fact that systems thinking is a very important higher-order thinking skill that should be fostered during classes. However, more attention has been dedicated to the different ways of the assessment of systems thinking skills, and less to their enhancement. Taking this into consideration, the goal of this study was not only to validate the systems thinking skills of secondary school students, but also to help students in the complex process of their development. With this goal, new instructional and assessment tools – systemic synthesis questions [SSynQs], were constructed, and an experiment with one experimental (E) and one control (C) group was conducted during organic chemistry classes. Namely, the instructional teaching/learning method for both E and C groups was the same in processing new contents, but different on classes for the revision of the selected organic chemistry contents. The results showed that students exposed to the new instructional method (E group) achieved higher performance scores on three different types of systems thinking than students from the C group, who were taught by the traditional method. The greatest difference between the groups was found in the most complex dimension of systems thinking construct – in the II level of procedural systems thinking. Along with this dimension, structural systems thinking and the I level of the procedural systems thinking were also observed.

Keywords: assessment tool; instructional tool; secondary chemistry education; systemic approach; higher-order thinking.

INTRODUCTION

In science education literature, the central goal of fostering students' higher-order thinking skills (HOTS),^{1–3} or higher-order cognitive skills (HOCS)^{4,5} has been widely accepted. Even though it is challenging to provide an exact defin-

* Corresponding author. E-mail: tamara.hrin@dh.uns.ac.rs

Serbian Chemical Society member.

doi: 10.2298/JSC160811097H



ition of HOTS, Resnick⁶ noted their non-algorithmic, multiple and complex nature. According to this notation, it is expected that HOTS include a wide range of different skills. Hence, these skills include problem solving,^{5,7} decision making,^{3,5,7} critical thinking,^{5,7} the ability to ask complex questions,^{1,3,7} generating argumentation^{1,3} and systems thinking.^{3,7}

In our opinion, a very detailed definition of systems thinking is that proposed by Salisbury.⁸ This author defined systems thinking as the ability to think about all the concepts and relations that exist within a system, in order to effectively structure these relations. Furthermore, Evagorou *et al.*⁹ simplified this definition, describing systems thinking as the ability of an individual to understand and interpret complex systems. In addition, system could be considered as an entity that exists and functions through the interactions of its parts.¹⁰ The term system is very broad, and appears in several domains: social systems, technological systems and natural systems. In recent years, it has been recognized that many phenomena around us are examples of complex natural systems (*e.g.*, ecosystems, hydro-cycle, metabolism, mixtures, organic compounds, *etc.*). For this reason, the understanding of complex systems and systems thinking has received growing attention in science education.^{9,11} Yet, the real problem was stated when researchers recognized students' limited understanding of complex systems.¹² In order to solve this problem, Gulyaev and Stonyer¹³ proposed building an integrated curricula and developing an instructional method based on a mapping model. However, it should be mentioned that science education literature offers just a few empirical studies about how students could become good systems thinkers,⁹ and how systems thinking skills could be promoted through teaching. Within this goal, computer simulations have often been used. Regarding elementary school students, Evagorou *et al.*⁹ examined the impact of simulation on the students' systems thinking skills in one science (biology) topic, *i.e.*, the ecosystem of a marsh. Additionally, the same instructional method was effective in developing the students' systems thinking skills in secondary school biology lessons, however only in combination with a specific lesson within the context of education for sustainability development (ESD).¹⁴ Along with simulations, a specifically designed computer game seemed to be effective in biology classes too.¹⁵ Furthermore, a slightly different approach to promoting the students' systems thinking skills was proposed by Assaraf and Orion¹⁰ in Earth science classes. These authors perceived the need for transforming the traditional classroom setting into specifically designed classes – a multidisciplinary-based study program. These classes should include a knowledge-integration activity, as well as an outdoor learning activity.

According to the previously mentioned studies, systems thinking is a very important HOTS that students should develop,^{9,16} but also very difficult to measure.¹⁶ Brandstädter *et al.*¹⁷ stated that, in addition to video analysis, the

most widely used tools for assessing students' systems thinking skills are questionnaires and interviews. This statement confirms a study by Assaraf and Orion,¹⁰ who used several different qualitative and quantitative research tools along with questionnaires and interviews. For example, by analyzing drawings and word associations of students, these two authors concluded that students possess an incomplete picture of the water cycle, as well as many misconceptions. On the other hand, observing the results of pre- and post-test analysis of concept maps, Assaraf and Orion¹⁰ noted the improvement of students in their ability to identify relations among concepts. Similar results were reported in a paper by Brandstädter *et al.*,¹⁷ who suggested computer and paper-pencil concept maps as adequate instruments for analyzing students' systems thinking. However, these authors took a step forward, observing two-dimensional way of systems thinking – structural and procedural systems thinking. The first dimension, or structural systems thinking, refers to analyzing the basic structure of the system – relevant concepts and their relationships. On the other hand, second dimension, or procedural systems thinking, refers to the understanding of dynamic and cyclic processes that occur within systems' elements and subsystems.¹⁷ The importance of this research was observed by Dauer *et al.*¹⁸ who applied box-and-arrow (similar to concept maps) as an instructional and assessment tool in biology courses. Even though the results of this study indicated that students improved their structural systems thinking ability, the authors did not examine the students' procedural systems thinking ability.

One of the latest papers is that by Vachliotis *et al.*,¹⁹ who examined secondary school students' systems thinking skills in an organic chemistry domain. For this purpose, the authors constructed and evaluated fill-in-the blank systemic assessment questions (SAQs). SAQs were introduced by Fahmy and Lagowski^{20,21} a few years later than their precursors – systemic diagrams.^{22,23} Systemic diagrams and SAQs take a central role within the context of systemic approach to teaching and learning chemistry (SATLC), as basic instructional, or/and assessment tools. According to Vachliotis *et al.*,²⁴ systemic diagrams and SAQs belong to the broader group of concept mapping techniques. However, while Novak *et al.*²⁵ highlighted hierarchical structure as the main characteristic of concept maps, systemic diagrams and SAQs do not possess hierarchical levels of concepts in their structures.²⁴ On the contrary, systemic diagrams and SAQs present closed conceptual structures,²¹ in which all relations between concepts are made explicit to the learners.²³ Namely, emphasizing cyclic interrelations among new concepts and those previously acquired, systemic diagrams and SAQs could provide more information to students than concept maps.

However, the main differences between systemic diagrams and SAQs should be mentioned. On the one hand, systemic diagrams represent all relevant concepts within a specific teaching unit, and therefore contain a large number of

concepts. On the other hand, SAQs contain a smaller number of concepts compared to systemic diagrams, as they were designed to effectively assess meaningful understanding and systems thinking after students have become familiar with a particular teaching theme.^{26,27} As was previously mentioned, in SAQs, all concepts are directly or indirectly interrelated forming a closed structure,²¹ which could follow various geometric shapes. Fahmy and Lagowski²⁰ proposed triangular, quadrilateral, pentagonal, hexagonal SAQs, depending on the number of concepts included in the structures. An example of a common form of pentagonal SAQ with five concepts is presented in Fig. 1. Panels A–E represent a selected collection of concepts (*e.g.* chemical compounds), while x , y , z , q , and w highlight the relation between concepts (*e.g.* reagents, light, catalyst, temperature, pressure).

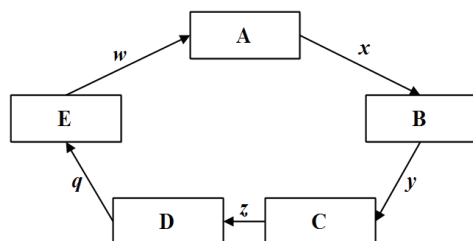


Fig. 1. Common form of pentagonal SAQ.

Besides their geometric shape, SAQs could also be distinguished by their type. Fahmy and Lagowski^{20,21} proposed various types of SAQs, *i.e.*, systemic multiple-choice questions (SMCQs), systemic true/false questions (STFQs), systemic sequencing questions (SSQs), systemic matching questions (SMQs), systemic synthesis questions (SSynQs), and systemic analyzing questions (SANQs). In the present study, SSynQs, as new instructional and assessment tools in systems thinking literature, were included. However, some changes were made to the original version. In Fahmy and Lagowski's²¹ version of SSynQs, all concepts are proposed in the statement of a particular task and students should develop (build) a diagram by positioning these concepts in the correct fields. Namely, they needed to determine the correct relations between concepts. On the other hand, the herein presented SSynQs were designed in a way that required recognizing, or identifying new concepts by observing the defined relations between them. Henceforth, students were faced with unfilled, or partially filled SSynQs (the initial concept was presented) in order to fill in the concepts that were missing.

It should be noted that fill-in-the-blank SAQs were previously characterized as valid and reliable tools for assessing students' meaningful and conceptual understanding of organic chemistry concepts by Vachliotis *et al.*^{19,24} In their latest paper, these authors¹⁹ included an additional variable in the research, *i.e.*, students' systems thinking skills. To the best of our knowledge, this is the only

study that considered students' systems thinking skills in the chemistry educational process. However, this study¹⁹ only considered the assessment of students' systems thinking skills, and it is still unknown how these skills could be promoted in the classroom.

METHODOLOGY

Study objective

The main objective of this study was to examine the impact of SSynQs on the learning gains of secondary school students. Such learning gains were observed through an assessment of experimental and control groups students' systems thinking skills in the domain of organic chemistry. From the defined research objective, two research questions arise:

1. Which test questions (conventional and/or SSynQs) could effectively assess the ability of secondary school students for structural and procedural systems thinking?
2. Which type of instructional method (traditional or the systemic approach by applying SSynQs) could more effectively provide an enhancement of students' structural and procedural systems thinking skills?

Sample

The study sample comprised 119 high school students (61 males and 58 females; four classes) of mixed ability and socio-economic status. The experiment was conducted during the second semester of the 2012/2013 school year, and focused on third-grade students (11th grade in the Anglo-American educational system; 17–18-year-old) who attended the same study stream: science and mathematics, in one high school in an urban region in Serbia. The participants of the study also included two chemistry teachers who have a Master's degree in Teaching Chemistry, and similar experience working with secondary school students (10 years of experience in teaching Chemistry). Each of the teachers taught two classes. In order to achieve the aim of this study, a quasi-experimental research design involving non-random assignment to two parallel groups, one experimental (E) and one control (C) was conducted. It should be clarified that the instructional method for both E and C groups was the same in the processing new organic chemistry contents, but different in revision classes. Namely, the E group students were treated by applying SSynQs, while C group was treated by applying the traditional method of teaching and learning the same material as the E group, however they did not receive any training in the systemic approach.

Such a research design with post-test only is involved in many studies in science education literature.²⁶⁻²⁹ Students were allocated to the different instructional conditions according to the average chemistry grades achieved at the end of previous school years (first and second school years). The average chemistry grade was 4.38 ($SD = 0.63$) in the E group and 4.42 ($SD = 0.70$) in the C group, in a five-point grading system. Group equalization based on average grades instead of initial testing was chosen to avoid the Hawthorne effect.^{30,31} Since the collected data was not normally distributed (Shapiro–Wilk test: E group: $W = 0.83$, $p = 0.000$; C group: $W = 0.79$, $p = 0.000$), application of the non-parametric Mann–Whitney test was chosen to compare the medians of the E and C groups. The obtained results showed that there was no statistically significant difference between E (mean rank, $MR = 58.77$, sum of ranks, $SR = 3820.00$) and C ($MR = 61.48$, $SR = 3320.00$) groups before the conducted research; for $U = 1675.00$, the p value was greater than 0.05 ($p = 0.653$). Thus, the formed E group contained 65 students, while the C group contained 54 students.

Context of the study

First, it should be mentioned that prior to the present study, a pilot study was performed with two groups (one E and one C) and one teacher. The design of the pilot study was similar to that of the present study, and it was conducted in the first semester of the 2012/2013 school year, with smaller sample of participants (two classes), who followed the organic chemistry course: "Hydrocarbons and their halogen derivatives". The main objective of the pilot study was to investigate initially the efficiency of SSynQs as assessment and instructional tools, examining the students' performance because of their meaningful learning,²⁶ as well as perceived mental effort.²⁷

In the present study (second semester of the 2012/2013 school year) students followed the organic chemistry course: "Organic compounds with oxygen". One teaching theme "Carboxylic acids and their derivatives" was chosen as the material for this experiment, which was processed in five school classes (one class lasted 45 minutes). In accordance with the curriculum regulations of the third grade of the science-oriented and general high school, as well as with the recommended textbook, the following contents were processed: 1) general properties of carboxylic acids, 2) classification of carboxylic acids, 3) preparation of carboxylic acids and their physical properties, 4) chemical properties of carboxylic acids and 5) the function and application of carboxylic acids. The main objective of the phase of the conducted experiment was to familiarize students with the basic features of the aforementioned teaching units. Namely, the transfer of information and accumulation of students' knowledge were realised through the teacher's lectures, discussion sections, and laboratory work (traditional method of teaching). During this period, all students were treated the same, relying on the recommended textbook and course material.

After the completion of this phase, the students were divided into E group (two classes) and C group (two classes). Hence, the E group students were taught by one chemistry teacher, while C group students were thought by the other chemistry teacher. One teacher was chosen for both E group classes for of several reasons. Firstly, the selected teacher has already participated in the prior study, and as such, was experienced in the systemic approach. Namely, the teacher was familiar with basic features of systemic approach and its applications in organic chemistry. Apart from this, Özmen *et al.*³² recognized that a teacher could have "experimenter bias", which could be explained by the fact that the teacher could be more enthusiastic about the experiment, and thus be more interested in the E than in the C group. For this reason, an experiment with the random assignment of the teachers to one E class and one C class was not chosen. Additionally, it is important to note that one author of this paper, who was a research assistant at the Faculty of Sciences, attended all the classes of both teachers in order to provide the same teaching quality.

The new teaching and learning environment

It should be mentioned that the revision of the selected organic chemistry contents (phase 2 of the study) was conducted during two school classes (90 min) in both E and C groups, so the length of time to study and revise was the same in both groups. For the purpose of the experiment, learning sheets with five unfilled and partially filled SSynQs were prepared, and presented to the teacher prior to their implementation in the classes. The SSynQs consisted of four to eight main concepts, and one example of the SSynQs with six concepts (main fields) is presented in Fig. 2.

Before receiving learning sheets, the students were introduced to the operational definitions of terms applied in SSynQs, *i.e.*, concepts, main relations, and cross-links. Additionally, before solving each SSynQ, they should read the instruction (presented in a text

above the SSynQ). By this, they could perceive the important relations between undefined concepts. Furthermore, the students were encouraged to participate actively in the classes, answering the teacher's questions and filling in the empty fields in the SSynQs with structural formulas and IUPAC-names of the required organic compounds. In this process, the teacher used a Power-Point presentation to ensure that all the students could see the correct answer at the end, which was presented *via* a video projector. The main goal of the experimental classes was to achieve a global view of selected chemical contents, connecting one class of organic compounds with others, previously learnt (*e.g.*, carboxylic acids and their salts with alkanes and alkyl halides; Fig. 2).

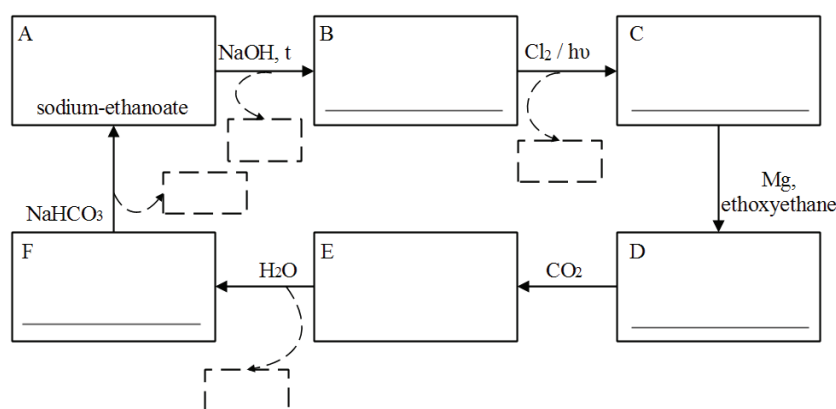


Fig. 2. Example of an SSynQ with six concepts that was applied in the experiment.

On the other hand, during phase 2, C group students continued with the application of the traditional method of teaching and learning of organic chemistry. Thus, students in the C group revised the same material as the E group, but they did not receive any training in the systemic approach. Namely, under the controlled condition (one author was present) they solved specifically constructed conventional tasks (*e.g.* multiple-choice questions, completion type questions). According to this, they receive instruction on the main systems thinking language: “links” that represent the influence of one concept on others. More precisely, such links highlighted the conditions (temperature, pressure, catalyst) and direction in which one specific concept could be transformed to the other. After completion of each task, the students received feedback about their response to a particular problem (task).

Research instrument

The instrument for assessing students' systems thinking skills (a specifically designed test of knowledge) consisted of 10 tasks. These tasks were divided into two main categories: SSynQs and conventional questions, which were previously used as valid and reliable tools for assessing systems thinking.^{14,15,17} By application of additional assessment tools along with SSynQs, a disadvantage position of the C group students was prevented.

Conventional questions were designed following several question types: matching question (Q1), completion type questions (Q2, Q10II), multiple-choice questions (Q3I, Q5, Q7, Q10I), and open-response questions (Q6, Q9). Additionally, SSynQs were presented by the numbers Q4 and Q8. For example, one open-response question (Q6) was defined as: “After finishing reaction between an aliphatic compound A with a molecular formula $C_5H_{10}O$, and

potassium dichromate solution acidified with dilute sulphuric acid, the orange solution turns green. The obtained substance B reacts with the sodium bicarbonate, forming a gas as one of the reaction products, which does not help burning (lighted wooden splint is extinguished). Draw structures and write IUPAC names for compounds A and B, if both compounds contain one chiral centre.” On the other hand, an SSynQ (Q4, Fig. 3) was defined as: “Write the structure of ethanoic acid in field A, and then, in the field B, write the structure and the IUPAC name of the compound that can be obtained by the reaction of ethanoic acid and calcium hydroxide. Compound B is further subjected to thermal decomposition to produce compound C. Compound C can be oxidized by hot nitric acid yielding two products: ethanoic acid and compound D. Compound D is reduced with lithium aluminium hydride to give compound E, which further reacts with compound F (obtained by a reaction of ethanoic acid and phosphorus trichloride) giving the final product G.” The correct answer for Q4 is presented in Fig. 3.

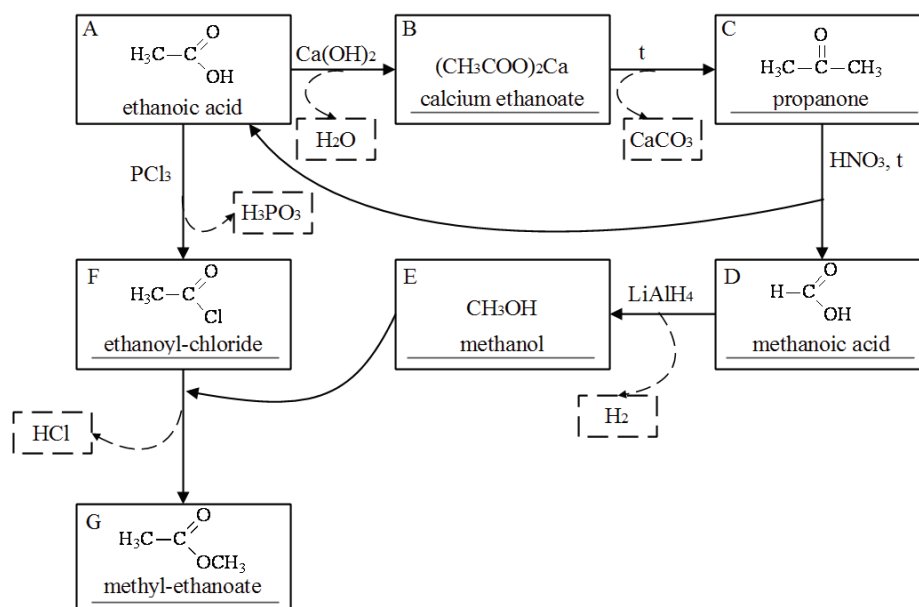


Fig. 3. Solved test question Q4 presented as an SSynQ.

Scoring of the performance on the test questions was performed based on the nature of the questions. Namely, multiple-choice questions were scored with one point, while in completion, open-response, matching questions, and [SSynQs], each individual requirement was scored with 0.5 points. In an SSynQ, 0.5 points were scored for each concept presented in the right field in a diagram by correct structural formula or IUPAC name, or with one point if such concept was presented by both the correct structural formula and IUPAC-name. The maximum performance score in the test was 28 points.

Quality assurance of the applied instrument. The credibility of the applied instrument was assured by considering the test validity and reliability. The content validity as a pre-test assurance parameter was estimated through the work of an expert team, composed of two university professors and two research assistants in the field of Chemistry Teaching, and one

high school chemistry teacher. They concluded that the test was valid as it was devised following the curriculum regulations of the selected chemistry course, as well as the recommended textbook. Additionally, the clarity and meaningfulness of the test question requirements, the variety of question types, the length of sentences and the applied terminology were considered.

Validity as a post-test assurance parameter was estimated considering the statistical results of the applied test, in accordance with Lopez *et al.*³³ In this analysis, the relation (Pearson's correlation coefficient) between SSynQs scores and performance scores on conventional questions was considered. After conducting a correlation analysis of the obtained results, the Pearson's coefficient of 0.774 ($p < 0.05$) showed, according to Dunn,³⁴ a strong positive, statistically significant correlation between the students' performance on [SSynQs] and conventional questions. This parameter confirmed the validity of the test.

The reliability was estimated by two different methods: *i*) internal consistency and *ii*) the split-half method. The reliability of internal consistency was calculated using the Cronbach α coefficient, which was found to be 0.925 (Table I). Additionally, the split-half method took into consideration several different coefficients (*e.g.* Cronbach's α for each part, the Spearman–Brown coefficients, *etc.*, Table I), which ranged from 0.855 to 0.922. The results of both methods showed that the applied test has excellent reliability.

TABLE I. Reliability calculations for the applied instrument

Internal consistency	Cronbach's α	0.925
Split-half method	Cronbach's α for part 1	0.865
	Cronbach's α for part 2	0.861
	Correlation between parts	0.855
	Spearman–Brown coefficient for equal lengths	0.922
	Spearman–Brown coefficient for unequal lengths	0.922
	Guttman split-half coefficient	0.921

RESULTS AND DISCUSSION

Assessing students' systems thinking skills: analysis of the proposed model

In order to answer the first research question, exploratory factor analysis (EFA) with maximum likelihood (ML) extraction was used. It should be mentioned that this study put 99 data into the factor analysis as not all students completed the test.

Before developing intervention and test items, it was hypothesized that the two-factor model (structural and procedural system thinking) proposed by Brandstädter *et al.*,¹⁷ would explain the students' responses to the set of 10 test questions. However, a chi-square measure of goodness-of-fit ($\chi^2 = 58.99$, $df = 26$, $p = 0.000$) rejected the hypothesis that two factors were adequate to explain the model. Thus, an alternative three-factor model was estimated. The Bartlett's test of sphericity ($\chi^2 = 682.10$, $df = 45$, $p = 0.000$; $p < 0.05$) and Kaiser–Mayer–Olkin measure of sampling adequacy ($KMO = 0.89$; $KMO > 0.50$) indicated the suitability of the EFA for the obtained data, and a chi-square measure of goodness-of-fit ($\chi^2 = 28.78$, $df = 18$, $p = 0.051$; $p > 0.05$) indicated that systems thinking

(ST) could be represented through the three factors. The three-factor model explained 76.97 % of the variance. The first factor was labelled as “structural ST”, the second factor as “I level of procedural ST”, and the third as “II level of procedural ST”. The different factors are explained below.

In order to determine the factor to which each test question is associated, the factor loadings were considered. The factor loadings were interpreted according to the Tabachnick and Fidell³⁵ model, by which the lowest acceptable value for factor loadings is 0.32. It should be mentioned that after conducting the EFA, all test questions were associated with the first factor, however, some of the questions were also associated with the second or the third factor (Table II). In the present case, it was not a surprising fact that all questions were loaded on the first factor as the test of knowledge was constructed to tap into the high school students’ ST skills, after applying selected instructional methods. Additionally, according to Brandstädter’s *et al.*¹⁷ definition of structural ST, students should have developed the ability to solve any of the ten test questions, identifying relevant elements (concepts) of the selected subsystem (“Carboxylic acids and their derivatives”) of one isolated system (“Organic compounds with oxygen”), and observe their interrelationships. The defined structural ST skills substantially correspond to the second identification step of the Vachliotis’s *et al.*¹⁹ ST construct.

TABLE II. The main factors and factor loadings for the test questions

Question number	First factor “Structural ST”	Second factor “I level of the procedural ST”	Third factor “II level of the procedural ST”
Q1	0.684	0.317	0.383
Q2	0.688	0.243	0.416
Q3	0.674	–	0.110
Q4	0.845	0.533	–
Q5	0.847	–0.531	–
Q6	0.610	0.519	–
Q7	0.603	–	0.269
Q8	0.786	–	0.486
Q9	0.503	0.129	–
Q10	0.610	–	0.243

Additionally, to solve questions Q1, Q2, Q4, Q6 and Q8 (Table II) the structural ST was not sufficient. For this reason, a one-factor model was not chosen, and these questions were associated with another factor. In literature, some authors noted that it is acceptable that some questions are loaded on two factors.^{19,24} For example, Q4 and Q6 were also associated with the second factor (“I level of the procedural ST”), while Q1, Q2 and Q8 were associated with the third factor (“II level of the procedural ST”). In order to explain the way in which some of the questions could evaluate lower or higher levels of students’ skills of

the procedural ST, the systems and subsystems that included these questions were examined. As can be seen from Table III, Q4 and Q6 have included one system (“Organic compounds with oxygen”) and several subsystems (Q4: “Carboxylic acids and their derivatives”, “Carbonyl compounds”, “Alcohols”; Q6: “Carboxylic acids and their derivatives”, “Carbonyl compounds”). Hence, in order to solve these questions, the students should interconnect several larger parts (subsystems), constituting a meaningful whole – conceptual system of interest. According to this, in our opinion, the I level of the procedural ST could actually be the third identification step of the ST construct developed by Vachliotis *et al.*¹⁹

TABLE III. Representation of associated systems and subsystems included in the test questions; System 1: organic compounds with oxygen

Question number	Subsystems of system 1	System 2	Subsystems of system 1
Q1	Carboxylic acids and their derivatives	Inorganic acids	Binary acids; oxyacids
Q2	Carboxylic acids and their derivatives	Hydrocarbons	Cyclic alkyl halides
Q3	Carboxylic acids and their derivatives		
Q4	Carboxylic acids and their derivatives; carbonyl compounds; alcohols		
Q5	Carboxylic acids and their derivatives		
Q6	Carboxylic acids and their derivatives; carbonyl compounds		
Q7	Carboxylic acids and their derivatives		
Q8	Carboxylic acids and their derivatives; alcohols	Hydrocarbons	Alkanes; alkenes; alkyl halides
Q9	Carboxylic acids and their derivatives		
Q10	Carboxylic acids and their derivatives		

On the other hand, the remaining three test questions, Q1, Q2 and Q8, showed not only interconnection of several subsystems of one isolated system, but also interconnection of more than one system, *i.e.*, two associated conceptual systems. Thus, Q1 included the additional system “Inorganic acids”, while Q2 and Q8 included “Hydrocarbons” (Table III). According to this, in order to solve the test questions associated with the third factor (II level of the procedural ST), the students should develop the ability to identify concepts and relationships of several subsystems of more than one system. In the paper by Vachliotis *et al.*,¹⁹ such ST skills were not mentioned.

Henceforth, after discussing the EFA results, all the test questions that were associated only with the first factor were considered as assessment tools for students' structural ST skills in the further analysis. Additionally, the test questions that were associated with the second or third factor along with the first factor were considered as assessment tools for the I or II level of students' procedural ST skills. The stability of the three-factor model was additionally confirmed by

calculating the Cronbach alpha of the separated subscales ($\alpha_1 = 0.843$; $\alpha_2 = 0.707$; $\alpha_3 = 0.871$).

Students' performance in the three types of STs

After determining which test questions could evaluate the students' ability of structural, or I and II level of procedural ST skills, the next aim of this research was to investigate the significance of the differences in the performance of the E and C group students in the three different types of STs. The basic statistical parameters obtained for these variables are summarized in Table IV, including the mean scores (M), the difference in the mean scores between the E and C groups (ΔM), the standard deviations (SD), the minima and maxima, and the ranges. The mean scores indicated that the E group achieved significantly higher scores in comparison to the C group, observing the performance in each of the three types of STs. From Table IV, it could be seen that the largest difference in performance between the groups was detected in the II level of the procedural ST, while the lowest difference appeared in the category of the structural ST. In particular, it was found that implementation of SSynQs as instructional tools increased the students' structural ST ability by 27.81 percentage points, the I level of the procedural ST ability by 30.74 percentage points, and the II level of the procedural ST by 33.21 percentage points. These differences are examined and discussed in detail in the following section.

TABLE IV. Descriptive statistics for the students' performance in the three different types of STs

Parameter	Performance in the structural ST		Performance in the I level of the procedural ST		Performance in the II level of the procedural ST	
	E	C	E	C	E	C
$M / \%$	86.25	58.43	84.44	53.70	58.17	24.96
$\Delta M / \%$	27.81		30.74		33.21	
SD	14.71	34.93	15.54	37.68	26.40	24.09
Minimum	50.00	0.00	33.33	0.00	12.50	0.00
Maximum	100.00	100.00	100.00	100.00	100.00	100.00
Range	50.00	100.00	66.67	100.00	87.50	100.00

In order to examine the statistically significant differences in the final examination scores between the two groups, the Mann–Whitney test with two independent samples (level of significance at $p < 0.05$), was applied. The statistical calculations were realised by application of the Nonparametric Tests from the statistical package IBM SPSS Statistics, version 19. For this analysis, the students' performances in the three types of STs were defined as dependent variables. It should be noted that the Mann–Whitney test was used after determining that the collected data did not satisfy the requirements of a normal distribution, which was confirmed using then Shapiro–Wilk test, as well as observing the

values for the standardized skewness and kurtosis. While the standardized kurtosis was within the +3 and -3 criterion³⁵ for all the examined situations, the standardized skewness did not meet an acceptable range in several of the situations. The values of standardized skewness for the performance of the E group students on structural ST ($ss = -3.15$) and on the I level of the procedural ST ($ss = -4.71$) highlighted the fact that the students' performances were significantly shifted towards higher scores. On the other hand, the same parameter for the performance of the C group students on the II level of procedural ST ($ss = 3.04$) indicated the presence of too many low scores. In addition, a heterogeneity of the variance between the groups was also detected (Levene's test: $F(5, 291) = 31.15, p = 0.000, p < 0.05$).

The results of the applied Mann-Whitney test, presented in Table V, showed that the E group had higher mean ranks than the C group for each of the three dependent variables. The assumption that the group with the highest mean ranks should have a greater number of higher scores was further confirmed with p -values and z -values. The p -values resulting from this test were lower than 0.05, while the obtained z -values were lower than the critical value, $z_c = -1.96$ (Corder and Foreman³⁶), observing the students' performance on the structural ST ($z = -4.31, p = 0.000$), on the I level of the procedural ST ($z = -4.35, p = 0.000$), and on the II level of the procedural ST ($z = -5.79, p = 0.000$). These parameters showed that the differences in the performance on the three types of ST of the E group, with respect to the C group, were statistically significant. In order to provide a clear description of the size of the observed statistically significant effects, using the z -values the effect size index (r) was calculated for all dependent variables, following the formula:³⁷

$$r = z / \sqrt{N} \quad (1)$$

Henceforth, using the r benchmarks provided by Cohen,³⁸ the effect sizes were evaluated. While medium effect sizes were determined for the performance in the structural ST ($r = -0.43$), and the I level of the procedural ST ($r = -0.44$), the students' performance in the II level of the procedural ST was characterized by a large effect size ($r = -0.58$).

TABLE V. Results of the Mann-Whitney test for students' performance in different types of STs; $p = 0.000$

Performance	Group	Mean rank	Sum of ranks	U	z
Structural ST	E (N=48)	62.74	3011.50	612.50	-4.31
	C (N=51)	38.01	1938.50		
I level of the procedural ST	E (N=48)	62.94	3021.00	603.00	-4.35
	C (N=51)	37.82	1929.00		
II level of the procedural ST	E (N=48)	67.21	3021.00	398.00	-5.79
	C (N=51)	33.80	1929.00		

From the results obtained in the Mann–Whitney test, it could be stated that students who applied SSynQs during learning the selected organic chemistry topic (“Carboxylic acids and their derivatives”) outperformed the students from the C group. The E group students managed to develop procedural as well as structural ST ability in a much more effective way than the students who were taught by the traditional method. Apart from this, the calculated effect sizes indicated that the most notable difference between the groups appeared in the highest level of ST. This could be explained by the fact that instruction *via* SSynQs brought students to a level in which they could not only identify the initial concepts (organic compounds) and simple relations, but also effectively “transform” such concepts within the selected system. According to Assaraf and Orion,¹⁰ this implies improved ability of students to include correct concepts into cycles (subsystems), identifying interrelations among them, and organize them into a complex system. However, the important question that follows is which characteristics of the applied instructional method provided such benefits, in comparison to the traditional instructional method. Perhaps, the right answer lies in the structure characteristics of [SSynQs]. Observing the structure and design of [SSynQs], these instructional tools could be characterized as completion of problems. Van Merriënboer and Sweller³⁹ defined completion problems within cognitive load theory (CLT) as problems (tasks) presented in a more or less completed (partially completed) form, in which students need to find the remaining concepts for themselves. In their study, Mihalca *et al.*⁴⁰ found that completion of problems, as instructional tools, stimulate students to process the solution more deeply, and enables them to construct complex cognitive schemas. According to Van Merriënboer *et al.*,⁴¹ “schemas provide a bridge between what learners already know and what they need to know to perform the learning task”. Finally, the ability to make such relations and networks within the selected system is the key characteristic of good system thinkers, and is highly relevant and appreciated in science education.

CONCLUSION AND IMPLICATIONS

The first aim of this study was to examine the effectiveness of [SSynQs], as new evaluation tools in systems thinking literature, in assessing the ST skills of secondary school students. Considering the paper by Brandstädter *et al.*,¹⁷ it was expected that a two-dimensional model of ST, structural and procedural STs, would be found and then analyzed. Contrary to expectations, a three-dimensional model was found where the [SSynQs] were effective in evaluating the I and II level of the procedural ST skills of students. These findings could be considered as valuable for future research, in which some other types of SAQs should be constructed and examined as tools for assessing different aspects of systems thinking construct.

With respect to the second suggested application of [SSynQs], the obtained results created a clear picture of [SSynQs] as effective instructional tools. Students from the E group managed to achieve higher performance scores on each of the three different types of STs, compared to the students from the C group. In addition, the highest difference between groups was found in the most complex dimension of ST – the II level of the procedural ST. These findings are in agreement with those of Hung,¹⁶ who stated that, in order to develop deep understanding of a selected complex system, students should develop the ability to integrate fully their knowledge. In this point, the traditional instructional method is usually not convenient.

Significance of the study. This study is significant in the area of science (chemistry) education because, to the best of our knowledge, this is the first paper that highlights a good way of how students' ST skills could be promoted in chemistry classes. Additionally, [SSynQs] were introduced and examined for the first time as new instructional and evaluation tools in the area of ST. By this, the results obtained in this study contribute to the growing body of both theory and empirical practice in ST literature, in the field of science education.

Limitations of the study. Limitations of this study could be observed from a methodological standpoint, since this study included students from only one high school study stream: science and mathematics. Additionally, only organic chemistry contents were investigated.

Implications for future research. Future research with another sample (e.g., another high school study stream, or another secondary school profile) would assist to confirm the results, or to obtain new ones. In addition, a more extensive coverage of organic chemistry material should have been tested. Apart from that, this research could be applied to different chemistry domains, such as general, inorganic chemistry, or biochemistry. However, the main future research attention should be focused on analyzing the possible gender differences, as well as differences in students' cognitive load (invested mental effort) in the defined dimensions of ST. By that, the connection between systems thinking constructs and the cognitive load theory would be posted. Such results have not been yet presented in science education literature.

Acknowledgement. This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia under Grant No. 179010.

ИЗВОД

ПРОЦЕНА И ПОБОЉШАЊЕ УЧЕНИЧКИХ ВЕШТИНА СИСТЕМСКОГ МИШЉЕЊА ПРИМЕНОМ СИСТЕМИЧКИХ ЗАДАТАКА СИНТЕЗЕ У ОРГАНСКОЈ ХЕМИЈИ

ТАМАРА Н. ХРИН, ДУШИЦА Д. МИЛЕНКОВИЋ, МИРЈАНА Д. СЕГЕДИНАЦ И САША ХОРВАТ

Природно–математички факултет, Универзитет у Новом Сагу, Трi Д. Обрадовића 3, 21000 Нови Сад

Многа истраживања у пољу образовања у природним наукама истичу чињеницу о важности системског мишљења, као вештине мишљења вишег реда које треба неговати у наставном процесу. Међутим, већа пажња је посвећена различитим методама процене системског мишљења, а мења његовом побољшању. Узимајући ово у обзир, циљ истраживања био је не само да се процене вештине системског мишљења код ученика средњошколског узраста, већ да се помогне ученицима у комплексном процесу његовог развоја. Са тим циљем, конструисани су нови инструкциони алати и алати евалуације – системички задаци синтезе [SSynQs], након чега је спроведен експеримент на часовима органске хемије, са једном експерименталном (Е) и једном контролном (К) групом. Наиме, инструкциони метод учења и обучавања за Е и К групу је био исти код обраде новог градива, али се разликовао на часовима понављања одабраних садржаја органске хемије. Резултати су показали да ученици подвргнути новом инструкционом методу (Е група) остварују боља постигнућа на три различите димензије системског мишљења од ученика К групе који су обучавани традиционално. Највећа разлика се појавила у најкомплекснијој димензији конструкта системског мишљења – у II нивоу процедуралног системског мишљења. Поред ове димензије, структурално системско мишљење и I ниво процедуралног системског мишљења су такође размотрени.

(Примљено 11. августа, ревидирано 28. септембра, прихваћено 27. октобра 2016)

REFERENCES

1. S. Avargil, O. Herscovitz, Y. J. Dori, *J. Sci. Educ. Technol.* **21** (2012) 207
2. M. Barak, D. Ben-Chaim, U. Zoller, *Res. Sci. Educ.* **37** (2007) 353
3. Y. J. Dori, R. T. Tal, M. Tsaushu, *Sci. Educ.* **87** (2003) 767
4. G. Tsaparlis, U. Zoller, *Univ. Chem. Educ.* **7** (2003) 50
5. U. Zoller, *J. Chem. Educ.* **70** (1993) 195
6. L. B. Resnick, *Education and Learning to Think*, National Academy Press, Washington, DC, 1987, p. 2
7. C. Nakiboğlu, H. E. Yildirim, *Int. J. Sci. Math. Educ.* **9** (2011) 1047
8. D. F. Salisbury, *Five Technologies for Educational Change: Systems Thinking, Systems Design, Quality Science, Change Management, Industrial Technology*, Educational Technology Publications, Englewood Cliffs, NJ, 1996, p. 17
9. M. Evagorou, K. Korfiatis, C. Nicolaou, C. Constantinou, *Int. J. Sci. Educ.* **31** (2009) 655
10. O. B. Z. Assaraf, N. Orion, *J. Res. Sci. Teach.* **42** (2005) 518
11. S. A. Yoon, *Int. J. Sci. Educ.* **30** (2008) 1
12. O. B. Z. Assaraf, N. Orion, *J. Res. Sci. Teach.* **47** (2010) 540
13. S. A. Gulyaev, H. R. Stonyer, *Int. J. Sci. Educ.* **24** (2002) 753
14. W. Riess, C. Mischo, *Int. J. Sci. Educ.* **32** (2010) 705
15. C. Sommer, M. Lücken, *Nor. Stud. Sci. Educ.* **6** (2010) 125
16. W. Hung, *Brit. J. Educ. Technol.* **39** (2008) 1099
17. K. Brandstädter, U. Harms, J. Großschedl, *Int. J. Sci. Educ.* **34** (2012) 2147
18. J. T. Dauer, J. L. Momsen, E. Bray Speth, S. C. Makohon-Moore, T. M. Long, *J. Res. Sci. Teach.* **50** (2013) 639

19. T. Vachliotis, K. Salta, C. Tzougraki, *Res. Sci. Educ.* **44** (2014) 239
20. A. F. M. Fahmy, J. J. Lagowski, *Afr. J. Chem. Educ.* **2** (2012) 66
21. A. F. M. Fahmy, J. J. Lagowski, *Afr. J. Chem. Educ.* **4** (2014) 35
22. A. F. M. Fahmy, J. J. Lagowski, *Pure Appl. Chem.* **71** (1999) 859
23. A. F. M. Fahmy, J. J. Lagowski, *J. Chem. Educ.* **80** (2003) 1078
24. T. Vachliotis, K. Salta, P. Vasiliou, C. Tzougraki, *J. Chem. Educ.* **88** (2011) 337
25. J. D. Novak, D. B. Gowin, G. T. Johansen, *Sci. Educ.* **67** (1983) 625
26. T. Hrin, D. Milenković, M. Segedinac, *Int. J. Sci. Math. Educ.* **14** (2016) 805
27. T. Hrin, A. F. M. Fahmy, M. Segedinac, D. Milenković, *Res. Sci. Educ.* **46** (2016) 525
28. S. Y. Kwon, L. Cifuentes, *Educ. Technol. Soc.* **10** (2007) 269
29. J. V. Powell, U. G. Aeby, T. Carpenter-Aeby, *Comput. Educ.* **40** (2003) 183
30. G. H. Bracht, G. V. Glass, *Am. Educ. Res. J.* **5** (1968) 437
31. S. R. G. Jones, *Am. J. Sociol.* **98** (1992) 451
32. H. Özmen, G. Demircioğlu, R. K. Coll, *Int. J. Sci. Math. Educ.* **7** (2009) 1
33. E. Lopez, J. Kim, K. Nandagopal, N. Cardin, R. J. Shavelson, J. H. Penn, *Chem. Educ. Res. Pract.* **12** (2011) 131
34. D. S. Dunn, *Statistics and Data Analysis for the Behavioral Sciences*, McGraw-Hill, New York, 2001, p. 205
35. B. G. Tabachnick, L. S. Fidell, *Using Multivariate Statistics*, Pearson Education, Boston, MA, 1996, p. 607
36. G. W., Corder, D. I. Foreman, *Nonparametric Statistics for Non-statisticians: A Step-by-Step Approach*, Wiley, Hoboken, NJ, 2009, p. 57
37. C. O. Fritz, P. E. Morris, J. J. Richler, *J. Exp. Psychol.* **141** (2012) 2
38. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, New York, 1988, p. 75
39. J. J. G. Van Merriënboer, J. Sweller, *Educ. Psychol. Rev.* **17** (2005) 147
40. L. Mihalca, C. Mengelkamp, W. Schnotz, F. Paas, *Contemp. Educ. Psychol.* **41** (2015) 157
41. J. J. G. Van Merriënboer, P. A. Kirschner, L. Kester, *Educ. Psychol.* **38** (2003) 5.