International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

TEACHING STUDENTS TO SELF-ASSESS USING COGNITIVE STRUCTURE ANALYSIS: HELPING STUDENTS DETERMINE WHAT THEY DO AND DO NOT KNOW

Celeste Cynkin and John Leddo

MyEdMaster, LLC

DOI: 10.46609/IJSSER.2023.v08i09.040 URL: https://doi.org/10.46609/IJSSER.2023.v08i09.040

Received: 10 September 2023 / Accepted: 23 September 2023 / Published: 4 October 2023

ABSTRACT

In previous papers, we have reported an assessment technique called Cognitive Structure Analysis (CSA) that is designed to assess the concepts that people have, not just how well they can give correct answers to questions. Types of knowledge covered by CSA include facts, procedures, problem solving strategies and rationales (why things work the way they do). Experimental tests of CSA have showed high correlations between the assessments of student knowledge and how well students perform on problem solving tasks. The present paper explores whether students can be taught how to use CSA to self-assess the knowledge they have about a topic they have just been taught. 16 students attending a gifted and talented high school were initially given instruction on how to self-assess using CSA. They were then given a lesson in calculus and asked to self-assess their own knowledge. They were then given a problem solving test that required knowledge of the topic they were just taught. Self-assessment protocols were evaluated with items listed in the protocols assigned to one of five categories: knowledge believed by the student to be relevant but not actually relevant; knowledge that the student knew was relevant and that the student did know; knowledge that the student knew was relevant but the student knew s/he did not have; knowledge that the student knew was relevant and believed s/he had but actually gave the wrong information; knowledge that was actually necessary to solve the problems but the student did not mention at all. Results showed that students' self-assessments, on average, were 92% accurate, meaning of the knowledge required for problem solving, either students both knew the relevance of the knowledge and gave the correct information ("knowing what you know") or they recognized that the knowledge was relevant and that they lacked the knowledge ("knowing what you don't know"). Further, when comparing these five knowledge categories to the problem solving scores, only the two categories that demonstrated correct self-assessment ("knowing what you know" and "knowing what you don't know") showed statistically significant correlations. These results indicate that students could be

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

able to determine what knowledge they lack for effective problem solving and that any gaps in their knowledge that they cannot effectively self-assess may not impair learning.

Introduction

Assessment has long been an integral part of the education process. It is seen as the measurement of how much students have learned the content that they were taught. In both classroom settings and in standardized testing, "learned the content" is typically operationally defined in terms of the number of correct answers a student gives on test questions. Indeed, classical test theory, one of the major pillars of assessment methodology assumes that the total number of correctly-answered test items indicates the students level of knowledge (cf., de Ayala, 2009).

Over the years, a number of assessment frameworks have been utilized by teachers and educational organizations. Typically, these can be categorized by whether students are asked to select the correct answer from a set of answer choices or asked to construct their own answers to problems. There has been considerable debate over which category of method is better, with pros and cons attached to each. Multiple choice tests require students to select answers from several distracters. Multiple choice tests are widely used in standardized testing and in many classroom settings due to the ease of grading (Chaoui, 2011) and the fact that students often score higher on multiple choice tests than they do on constructive response tests as students can increase their scores through guessing (cf. Elbrink and Waits, 1970; O'Neil and Brown, 1997). However, such guessing is often cited by critics as a reason why multiple choice tests should not be used.

At the other end of the continuum are constructive tests, which require that students enter answers to questions rather than choose from different answer choices. Researchers find, when investigating math problem solving, that students are more likely to use guessing strategies when given multiple choice tests but are more likely to reason mathematically when given constructive tests (Herman et al., 1994), thus making the test more ecologically valid in measuring students' actual knowledge (Frary, 1985).

The challenge with the key assumption of classical test theory, that correct answers indicate learning and vice versa, is that this assumption may not be entirely true. A medical analogy works well here. Normally, if a person shows outward signs of illness, s/he is probably sick (although there could be non-medical reasons why a person can appear sick such as overexertion or lack of sleep). Similarly, a student who makes a lot of mistakes on a test probably has a lack of knowledge (unless, for example, s/he was distracted or sick during the test). However, a person can look healthy and still have an underlying illness. Similarly, a student may get correct

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

answers on a test and have knowledge deficiencies. They can be parroting facts or formulas that they do not really understand or guessing correctly on multiple choice exams (which is a major criticism of that testing format).

More importantly, the lack of correct answers does not inform the teacher as to what concepts need to be remediated. A doctor does not stop his/her diagnosis after observing symptoms. The doctor runs further tests to discover the cause of the symptoms, so that an appropriate remedy can be applied. Indeed, we would consider it medical malpractice for a doctor to treat only the symptoms and not the underlying causes of diseases. Similarly, an incorrect answer to a test question is a symptom that may indicate an underlying knowledge deficiency. We consider it to be educational malpractice to stop the assessment at that point without diagnosing the underlying knowledge deficiency that is causing that incorrect answer. Unless that cause is identified, how can the appropriate remedial instruction be given?

Previously, we have reported an assessment methodology called Cognitive Structure Analysis (CSA) that is designed to assess the underlying concepts a student has, so that when a student does make a mistake, the source of that mistake can be identified and remediated (Leddo et al., 2022; Ahmad and Leddo, 2023; Zhou and Leddo, 2023). CSA is based on decades of cognitive psychology research that have shown that people possess a variety of knowledge types, each of which is organized and used differently in problem solving. Because there are different types of knowledge that people have, our framework is an integration of several prominent and well-researched formalisms. These include: semantic nets, which organize factual information (Quillian, 1966); production rules, which organize concrete procedures (Newell and Simon, 1972); scripts, which are general goal-based problem solving strategies (Schank and Abelson, 1977; Schank, 1982); and mental models, which explain the causal principle behind concepts (de Kleer and Brown, 1981). Because our framework integrates these four knowledge types, it is called INKS for INtegrated Knowledge Structure.

The INKS framework is based on research by John Leddo (Leddo et al., 1990) that shows that true expertise in a subject area requires all four knowledge types. INKS also has implications for instruction. For example, in John Anderson's ACT-R framework, people initially learn factual/semantic knowledge that is later operationalized into procedures (Anderson, 1982). Research by Leddo takes this one step further showing that expert knowledge is organized around goals and plans (referred to in the literature as "scripts" – Schank and Abelson, 1977; Schank, 1982) and abstracted into causal principles (referred to in the literature as "mental models" – cf., de Kleer and Brown, 1981) that allow people to construct explanations and make predictions/innovations in novel situations.

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

In order to identify the root cause of the mistake, we use a query-based assessment framework called Cognitive Structure Analysis (CSA, Leddo et al., 1990). CSA incorporates principles from the INKS knowledge representation framework. In our recent research, assessments produced by CSA correlated .966 with problem solving performance in Algebra 1 (Leddo et al, 2022), .63 with problem solving performance in the scientific method (Ahmad and Leddo, 2023), and .80 with problem solving performance in pre-calculus (Zhou and Leddo, 2023).

As argued above, the value of CSA as an assessment tool is that it can provide teachers with a means of assessing what concepts students have and are lacking, so that appropriate remedial instruction can be provided. However, this line of reasoning presupposes that a teacher is the one who does the assessment and remediation. The present paper raises the research question of whether students can be taught to perform CSA on their own knowledge, and in doing so, can seek their own remedial instruction and improve learning outcomes. There are two benefits that teaching students to self-assess could have for the educational process.

The first is that students will not have to wait until they receive corrective feedback from teachers on what they do not understand. By self-assessing, students can be proactive and seek out what they diagnose as needed learning. Second, given the proliferation of online resources these days, students are increasingly engaged in self-directed learning. Self-assessment can be a crucial part of the self-directed learning process by enabling students to assess how well they have learned the material they seek to learn and to guide their further instruction.

The present paper explores whether students can learn to self-assess using the CSA technique. Since this is the first attempt we are making at teaching students to self-assess, we use gifted and talented (GT) high school students as the initial group to investigate. Our previous research in self-directed learning (Leddo et al., 2017; Nittala, Leddo and Nittala, 2022; Leddo and Kalwala, 2023) suggests that GT students are less dependent on teachers for high levels of learning than non-GT students are. Therefore, we expect that GT student would be most able to learn to self-assess. The testbed chosen in calculus. In the present investigation, we are concerned with two research questions: 1) how accurate are the self-assessments produced by students?; 2) how well do these self-assessments correlate with actual problem solving performance?

Methods

Participants

Participants were 16 tenth to twelfth grade students who attended Thomas Jefferson High School for Science and Technology, a magnet school for gifted and talented students. This grade range ensured that the participants had learned the concept of limits, but would have a wide range of experience when it comes to the concept of derivatives. This means there would be students with

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

no knowledge of derivatives, as well as students who have been practicing daily with concepts built upon derivatives.

Materials

A lesson on the concept of derivatives wascreated forParticipants, including topics such as tangent lines, instantaneous rate of change, and the definition of a derivative. Below is the link to the videos watched and the concepts taught.

https://docs.google.com/document/d/1y7OFBmz_E2QeE1atEPVe_7vUWLjlyhrOnD-IYlV8cCI/edit

An example of self-assessment was also shown to students to replicate with their new knowledge of derivatives. Below is the example:

Script for self-assessment

"I want to teach you how to assess your own knowledge that you have about a subject area. Let's do this by taking an example that you already know. Suppose you wanted to assess your own knowledge about solving 2-step equations of the form ax + b = c. An example of this type of problem is 2x + 3 = 15. If I want to be able to solve problems like these, I need four types of knowledge. These are facts, strategies, procedures and rationales. Facts are concepts you have that describe objects or elements. For example, for two step equations, I need to know what variables, constants, coefficients, equations, and expressions are. Strategies are general processes I would use to solve a problem. For two step equations, this would be reverse order of operations. Procedures are the specific steps that I would use in a strategy. So if I am using reverse order of operations, I need to know additive and multiplicative inverses. Finally, I need to know rationales which are the reasons why the strategies or the procedures work the way they do. For example, this could include things like the subtraction or the division property of equalitythat says that when you do the same operation to both sides of an equation, you preserve the value of the equation. You can think of facts as telling you "what", strategies and procedures as telling you "how" and rationales as telling you "why". With this in mind, this is how I might assess my own knowledge of solving two step equations.

For facts, I need to know what variables, constants, coefficients, equations and expressions are. A variable is an unknown quantity, usually represented by a letter. A constant is a specific number. A coefficient is a number that you multiply a variable by like 2x. An equation is an expression that is equal to another expression and the two expressions are joined by an equal sign. An expression is one or more terms that are combined by mathematical operations like addition, subtraction, multiplication and division.

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

For strategies, I need to know the reverse order of operations which is SADMEP. This stands for subtraction, addition, division, multiplication, exponents and parentheses. I know that I'm supposed to do these in order but I don't remember whether I'm supposed to do subtraction always before addition or just which one goes first. The same is true for division and multiplication.

For procedures, I need to know additive inverse and multiplicative inverse. The additive inverse is taking the number with the opposite sign as the constant and adding it to both sides of the equation. The multiplicative inverse is taking the inverse of the coefficient of the variable and multiplying both sides of the equation by it. However, if the coefficient is negative, I'm not sure if the multiplicative inverse is supposed to be negative as well.

For rationales, I believe the two rationales I need are the subtraction property of equality and the division property of equality. The subtraction property of equality says that if I subtract the same number from both sides, which is what I'm doing with the additive inverse, I preserve the equality. Similarly, the division property of equality says that if I divide both sides of the equation by the same number, which is what I'm doing with the multiplicative inverse, I preserve the equality. When I look over what I wrote, I see that I am good with my facts.

On my strategy, I'm not sure about the order of steps in reverse order of operations when it comes to subtraction and addition or multiplication and division, so I need to learn those. On procedures, I'm not sure what to do with multiplicative inverses when the coefficient is negative, so I need to learn that as well. For rationales, I think I'm OK. I don't think I have any missing facts/concepts that I left out that I should know or I didn't list any facts/concepts where I didn't know what they were. For the strategy, I believe I listed the correct strategy and parts of the strategy, but I wasn't sure about some of the ordering of steps in the strategy. For procedures, I was good on the additive inverse but had a question on carrying out the multiplicative inverse when the coefficient was negative. For rationales, I think I had all the rationales that were important and that I understood them as well. I don't think I left anything out."

The problem solving assessment administered to the students to gauge their knowledge is linked below:

https://docs.google.com/document/d/161dHrrqaTsE4USS7HhAK6RFCoYiPwEHRqnxJ-i_jL28/edit?usp=sharing

Procedure

After the participants were given a lesson on derivative concepts, they completed a self-assessment on their own. In the self-assessment, they were told to list their amount of knowledge on each topic in the four categories of fact, procedure, strategy, and rationale.

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

A problems solving test was administered to the participants to test their knowledge of derivatives. The test consisted of 8 free response questions which tested facts, procedures, strategies, and rationales.

Results

Participants' self-assessments and post-test scores were evaluated. In order to evaluate the self-assessments, a list of facts, procedures, strategies and rationales was created and then compared to the Participants' self-assessment protocols. A score of 1 was given for each correct test answer and for each self-assessment item that matched the list items described above. Scores of 1 were also given for each occurrence of errors and omissions, the categories of which are explained below.

The mean number of correct post-test answers and self-assessed items are shown in the table below.

Column A presents the mean number of questions that were answered correctly out of eight questions on the test.

All the other data and information that are presented in the table related to the self-assessment.

Column B presents the mean number of irrelevant information that was provided in the self-assessment. This can be thought of as a "false alarm," namely, the Participant thought the knowledge was relevant to the topic but was not.

Column C presents the mean number of items on the self-assessment that were also in the list. This can be thought of as "hits"or "knowing what you know."

Column D presents the mean number of items that Participants stated that they did not know and were actually on the list, meaning Participants should know them. This can be thought of as correctly "knowing what you don't know."

Column E presents the mean number of items that are on the list that Participants recognized that they needed to know but actually gave the wrong information. This corresponds to believing one knows something but does not know it.

Column F presents the mean number items on the list but not mentioned by the Participants in their self-assessments. This represents knowledge that Participants needed to have but did not realize they needed to have that knowledge.

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

Columns C, D, E and F total to 10 as there were 10 different pieces of information on the list of information on derivatives. Column B has no specific number limit since there is no limit to the number of items a person can false alarm to.

A	B	C	D	E	F
Number of problems correct	Things that they think are important but are not on the list	Thinks they know, and are on list	Thinks they don't know, and are on the list	Are on the list but they gave wrong answer	Are on the list but not mentioned
7.25	0.5	8.8125	0.375	0.0625	0.75

The first research question posed in the Introduction is the accuracy of the self-assessments. Technically, columns C and D represent accurate self-assessments and columns B, E and F represent inaccurate self-assessments. However, one can argue that column B, false alarms, is not a critical self-assessment issue since column B represents false alarms or things Participants think they need to know but are not on the list. In other words, this knowledge is irrelevant to proficient problem solving. The implication of column B knowledge is that Participants will seek remedial knowledge on information that is irrelevant and perhaps waste some time, or they will hold false knowledge about concepts that are irrelevant to their problem solving and, therefore, should have no impact.

Accordingly, the key areas of inaccurate self-assessment are shown in columns E and F. Of these, incorrect knowledge (having the wrong knowledge about something that is necessary to know) had very low occurrences with a mean of .0625. Complete misses (needing to know something but not knowing that one needs to know it) was higher at a mean of .75 items, but even this number is relatively low, especially compared to the mean number of correctly self-assessed items of 9.1875. Viewed another way, Participants' self-assessments were, on average, about 92% correct. A paired t-test that compared total accurate assessment items with inaccurate assessment items yielded the highly significant t vale of 17.07, df=15, $p < .0001$.

The second research question posed in the Introduction section is how well these self-assessments correlated with problem solving performance. In order to address this, the correlations between each self-assessment category from the table above and the number of correctly answered post-test scores were calculated. Of these, only two were significant. The correlation between correctly identified knowledge components (Column C) and problem-solving performance was .53, df = 14, $p < .04$. This suggests that the better Participants were at self-assessing what they do know, the better their performance was. The correlation between correctly identifying knowledge components Participants did not know and problem solving performance was -.58, df = 14, $p < .02$. This suggests that the more Participants realized what they did not know, the worse they did on problem solving. This suggests that Participants' ignorance did hurt their problem-solving performance, but also represented an opportunity for self-remediation.

All remaining correlations between self-assessment categories and problem-solving performance were highly insignificant. This is actually a positive finding. It suggests that the knowledge categories where Participants showed genuine ignorance of what they did and did not need to know and, therefore, were unlikely to seek any corrective instruction, did not impact their performance. Put another way, Participants paid no price for the areas of complete ignorance.

Discussion

The results of the present study were very encouraging in that they strongly suggest that students can use the CSA assessment technique to self-assess their subject matter knowledge. Key findings are that the self-assessments produced were highly accurate with an average accuracy of 92%. What is perhaps more important is that when a lack of knowledge did occur, Participants showed general awareness of the missing knowledge that negatively impacted their performance. Knowledge deficiencies and inaccuracies that Participants were not aware of did not impact performance. These results suggest that teaching students to self-assess using CSA could provide a valuable source of feedback on what students need to learn to boost their performance. Moreover, the results suggest the encouraging finding that students' potential limitations in their self-assessments (believing concepts are important when they are not, having the wrong information about concepts, or missing concepts entirely that they are not aware they are missing) have no impact on their performance.

Further research is suggested by the present study. First, the present study had only GT students as participants. This raises the question of how successful non-GT students would be in learning CSA as a self-assessment technique. Related to this, the present study focused on high school students. It is important to see whether younger students can learn to self-assess as well.

Second, the present study focused on calculus. It would be interesting to investigate how successful self-assessment training would be on other subjects. We have tested CSA as an assessment method (as opposed to a self-assessment method) on a few different subjects, and we have found that it is not equally effective on each subject (although it is statistically significantly

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

effective for each subject tested to date). CSA produces assessments that correlate .966 with problem-solving performance in Algebra 1 (Leddo et al., 2022), .80 with problem-solving performance in pre-calculus (Zhou and Leddo, 2023) and .63 with problem-solving performance in the scientific method (Ahmad and Leddo, 2023). The trend here is that the more routinized the subject matter, the greater the predictive power of CSA is in predicting problem-solving performance. The same trend may hold up using CSA as a self-assessment methodology.

Third, and perhaps most important, research needs to be conducted to investigate whether students can self-remediate based on the CSA-based self-assessments and whether such self-remediations can boost student achievement. If this occurs, then there is a potential paradigm shift possible in education.

The present results are promising for using CSA to help students self-assess their knowledge. While, historically, schools have relied on teachers to assess students and provide corrective instruction, this has been a challenge when teachers have classes with 20 or more students, each with potentially different needs. With a proliferation of educational content available to anyone these days, learning no longer requires the presence of a teacher. As a result, new tools are needed to support people in their learning process. Helping them to assess their own knowledge is a way to support people's learning and can potentially enable them to improve their learning outcomes.

References

Ahmad, M.&Leddo, J. (2023). The Effectiveness of Cognitive Structure Analysis in Assessing Students' Knowledge of the Scientific Method. International Journal of Social Science and Economic Research, 8(8), 2397-2410.

Anderson, J.R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369-405.

Chaoui, N (2011) "Finding Relationships Between Multiple-Choice Math Tests and Their Stem-Equivalent Constructed Responses". CGU Theses & Dissertations. Paper 21.

de Ayala, R. J. (2009). The theory and practice of item response theory. New York: The Guilford Press.

de Kleer, J. and Brown, J.S. (1981). Mental models of physical mechanisms and their acquisition. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Elbrink, L., & Waits, B. (Spring, 1970). A Statistical Analysis of MultipleChoice Examinations in Mathematics. The Two-Year College Mathematics Journal, 1(1), 25-29.

ISSN: 2455-8834

Volume:08, Issue:09 "September 2023"

Frary, R. (Spring, 1985). Multiple-Choice Versus Free-Response: A Simulation Study. Journal of Educational Measurement, 22, 21-31.

Herman, J. L., Klein, D. C., Heath, T. M., & Wakai, S. T. (1994). A first look: Are claims for alternative assessment holding up? (CSE Tech. Rep. No. 391). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.

Leddo J., Boddu B., Krishnamurthy S., Yuan K., & Chippala S. (2017). The effectiveness of selfdirected learning and teacher-led learning on gifted and talented vs. non-gifted and talented students. International Journal of Advanced Educational Research, 2(6):18-21.

Leddo, J., Cohen, M.S., O'Connor, M.F., Bresnick, T.A., and Marvin, F.F. (1990). Integrated knowledge elicitation and representation framework (Technical Report 90-3). Reston, VA: Decision Science Consortium, Inc.

Leddo, J. & Kalwala, S. (2023). The Effectiveness of Self-directed Learning and Teacher-led Learning of Advanced Subject Matter on Gifted and Talented Students. International Journal of Social Science and Economic Research, 8(9), 2819-2825.

Leddo, J., Li, S. & Zhang, Y. (2022). Cognitive Structure Analysis: A technique for assessing what students know, not just how they perform. International Journal of Social Science and Economic Research, 7(11), 3716-3726.

Newell, A. and Simon, H.A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice Hall.

Nittala, T., Leddo, J. & Nittala, H. (2022). The Effectiveness of Self-directed Learning Vs. Teacher-led Learning of Advanced Subject Matter on Gifted and Talented Vs. Non-gifted and Talented Students. International Journal of Social Science and Economic Research, 7(8), 2619-2627.

O'Neil Jr., H., & Brown, R. (1997). Differential Effects Of Question Formats In Math Assessment On Metacognition And Affect. Applied Measurement in Education, 331-351.

Quillian, M.R. (1966). *Semantic memory.* Camridge, MA: Bolt, Beranek and Newman.

Schank, R.C. and Abelson, R.P. (1977). *Scripts, Plans, Goals, and Understanding.* Hillsdale, NJ: Erlbaum.

Schank, R.C. (1982). *Dynamic Memory: A theory of learning in computers and people.* New York: Cambridge University Press.

Zhou, L.N. & Leddo, J. (2023). Cognitive Structure Analysis: Assessing Students' Knowledge of Precalculus. International Journal of Social Science and Economic Research, 8(9), 2826-2836.

