International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

COGNITIVE STRUCTURE ANALYSIS: A TECHNIQUE FOR ASSESSING WHAT STUDENTS KNOW, NOT JUST HOW THEY PERFORM

John Leddo, Shangzhi Li and Yujie Zhang

MyEdMaster, LLC

John Leddo is the director of research at MyEdMaster.

DOI: 10.46609/IJSSER.2022.v07i11.011 URL: https://doi.org/10.46609/IJSSER.2022.v07i11.011

Received: 15 November 2022 / Accepted: 26 November 2022 / Published: 30 November 2022

ABSTRACT

Assessment has been a key part of education, playing the role of determining how much students have learned. Traditionally, assessments have focused on whether students give the correct answer to problems, implying that that the number of correctly-answered test items is a valid measure of how much students know. Unfortunately, the focus on correct answers has also resulted in neglecting the potential ability of assessments to provide diagnostic feedback to educators as to what concepts students have mastered and where the gaps in their knowledge are, thus potentially informing the day-to-day teaching process. The present paper describes an assessment technique called Cognitive Structure Analysis that is derived from John Leddo's integrated knowledge structure framework (Leddo et al., 1990) that combines several prominent knowledge representation frameworks in cognitive psychology. Using a Google Form, students were queried on four types of knowledge considered the basis of mastery of Algebra 1 concepts: factual, procedural, strategic, and rationale. From students' responses to these queries, measures of each type of knowledge and a combined knowledge score were created. Students were also given problems to solve. Correlations between each knowledge component score and problem-solving performance were high and the correlation between overall CSA-assessed knowledge and problem-solving performance was a near-perfect .966. Results suggest that CSA can be both easily implemented and highly diagnostic of student learning needs. Future research can investigate CSA's robustness across other subjects and whether incorporating CSA as part of day-to-day classroom instruction can lead to higher student achievement.

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

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

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 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

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

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?

The present paper reports 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. 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.

We note that the National Council of Teachers of Mathematics (2000) has developed a taxonomy of strands necessary for students to be considered mathematically proficient that uses similar terminology: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning. In many ways, the strands of conceptual, procedural and strategic do correspond to our own. The key difference is that the National Council of Teachers of Mathematics frames these strands in terms of desired skills/behavioral outcomes whereas we conceptualize these in terms of the specific knowledge needed to achieve those outcomes.

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:07, Issue:11 "November 2022"

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. CSA is chosen because previous research shows that there is a strong correlation between user knowledge as assessed by CSA and performance in practical problem solving. In one previous research project, we found that the using an automated multiple-choice CSA system to assess student learning produced measures of knowledge that correlated .88 with student problem solving performance and measures of change of knowledge as a result of instruction that correlated .78 with change in performance from pre-test to post-test. Moreover, at-risk students taught based on the needs diagnosed using CSA performed at a mainstream level three grades higher than their own after a 25-hour tutoring program in science (Leddo and Sak, 1994). In two other projects, assessments produced using the CSA methodology produced assessments of student learning agreed with teachers' assessments approximately 95%-97% of the time, which was statistically equal to teachers' assessments with each other (Leddo et al., 1998, Liang and Leddo, 2021).

In the present project, topics in Algebra 1 were analyzed for the facts (semantic knowledge), strategies (scripts), procedures (production rules) and rationales behind the concepts (mental models). Open-ended questions were then constructed to see if students possessed each of these knowledge components. Students' answers to the questions were analyzed for their correctness. This allows us to obtain knowledge scores for each of the four types of knowledge that are considered necessary to master an algebra topic as well as an overall knowledge score. These scores are then correlated with students' performance in problem solving to see how well they predict that performance.

Method

Participants

Participants were 20 middle and high school students. In order to get a range of knowledge levels across participants, students were recruited with a wide range of mathematical experience. At minimum, students were at least already taking Algebra 1, but may not have completed it. At the higher end, some students were already enrolled in calculus.

Materials

One mathematics assignment was created covering topics from Algebra 1. Topics and questions were chosen from online math textbooks created by leading American publishers.

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

The subjects for the assignments were: linear equations with variables on both sides, linear equations with unknown coefficients, compound inequalities, solving equations with square roots, and quadratic formulas.

The assessment was composed of three sections. The first section included fact-based problems to examine students' understanding of the material's definitions. Fact-based questions for section one included:

Facts-Based Questions:

"What is a variable? How is it represented in a problem?"

"What is a square root?"

"What do you need to pay attention to when solving an inequality question? What are some of the differences between inequality and equation problems?"

"What is a coefficient? Where do you find the coefficient in a problem?"

"What is a radical?"

"How do you identify a coefficient in a linear equation with unknown coefficients?"

"What is the standard form of the quadratic equation?"

"What is the difference between "and" and "or" when solving inequality problems?"

"What is a constant? Where do you find the constant in a problem?"

"What is the relationship between a radical and square root? How do you identify a square root in a problem?"

Procedural Questions:

"How do you combine like terms?"

"How do you remove the constants to one side of the equation and variables to the other side of the equation?"

"How to isolate the variable in a linear equation with variables on both sides?"

"If you subtract anumber on one side of the equation, how do you perform it on the other side of the equation?"

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

"How do you perform additive inverse and multiplicative inverse when a negative number is involved? Are there any changes to the signs?"

"How do you represent your set of answers on the graph when solving compound inequalities questions?"

"How do you determine what is your final answer in "and" or "or" condition when solving compound inequalities questions?"

"How do you isolate the radical term in an equation?"

Rationale-Based Questions:

"Why do you need to keep both sides equal in previous procedures? Why is it important to keep the equation balanced?"

"Why is it important to keep variables on one side of the equation, and keep constant and unknown coefficients on the other side?"

"Why do you need to express the final result based on two solutions when solving compound inequalities questions?"

"Under what conditions will you only acquire one final result, and under what conditions will you have a set of results when solving compound inequalities questions?"

"Why do you need to square both sides of the equation to solve a radical problem?"

"Why should you simplify the equation after eliminating the radical?"

The second section was the strategy section, where students are asked only to write out the general strategies they use for different topics when solving Algebra 1 questions. The third section was a test section where students are asked to solve mathematical problems chosen from linear equations with variables on both sides, linear equations with unknown coefficients, compound inequalities, solving equations with square roots, and quadratic formulas topics. In the third section, students did not have access to what they have written before and were asked to demonstrate all procedures when solving mathematical problems.

The Google Form is shown below:

https://docs.google.com/forms/d/1GGamRwF89xZdJ1K8CucUAfWh9lhEdBJWAAYW_WF3dec/edit

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

Procedure

The materials were administered in the form of a Google Forms survey. Participants were given links to the survey and asked to fill out the survey. They were given as much time as needed. No calculators or outside resources were allowed. Participants were supervised to prevent any use of outside resources.

Results

Students' written results were analyzed based on definitions provided in online math textbooks created by leading American publishers or educational learning sites such as Khan Academy. All students' results were analyzed using the same material and the same standard.

The written section has 24 questions, the math strategy section had five questions, and the math problem-solving section had 24 questions, making the entire questionnaire worth 53 points in total. Each question was worth one point. When analyzing written questions, students' answers were evaluated by whether they had demonstrated a thorough understanding of the question instead of the word count. Students' answers that demonstrated a similar meaning to the official answers were given full credit (1 point). Students' answers that were partly correlated with the official answers were given partial credit (0.5). Students' answers that were completely different from official answers were given no credit (0).

In the strategy section, students were given five mathematical problems and were asked to write down all the strategies needed to solve each problem. Students' results were analyzed based on whether they had fully demonstrated all the strategies needed to use to solve the problem. Students' responses were compared to answer keys provided by math-major undergraduate students from New York University. If students wrote all strategies needed to solve the problem, they received full credit (1 point). If students wrote parts of strategies needed to solve the problem, partial credit was given (0.5). If students did not write any strategy, no credit was given (0). There were five students, however, who gave a step-by-step solution to the problems rather than describe their actual problem-solving strategies. We assumed that this represented a misreading of the instructions rather than a lack of strategic knowledge on their part. Accordingly, those five students were excluded from the analyses that looked at the relationship between strategic knowledge and problem-solving performance and the relationship between overall knowledge and problem-solving performance. Those students were included in the analyses of the relationship between factual, procedural and rationale knowledge and problem-solving performance since those analyses did not use measures of strategic knowledge.

In the test section, students were given 24 mathematical problems and were asked to solve each question and demonstrate all steps. Questions were chosen from five topics: linear equations with

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

variables on both sides, linear equations with unknown coefficients, compound inequalities, solve equations with square roots, and quadratic formulas. Students' answers were graded based on answer keys provided by math major undergraduate students from New York University. Students receive full credit (1 point) if they demonstrated all steps and calculated the correct result. Partial credit (0.5) was given to students who made errors in signs in the final answers but still acquired the correct numerical result. Students received 0 points if they did not get the correct answer.

In order to determine how well the INKS model could be used to model students' algebra knowledge and how well the CSA technique could be used to assess how much of that knowledge students have in a way that predicts their problems solving performance, the knowledge component scores were correlated with problem solving scores. The results of this analysis showed a correlation between total INKS knowledge as assessed by CSA and problem-solving performance of .966, df =13, p < .001.

The next step was to look at individual components of the CSA framework: factual, procedural, strategic, rationale and see how well they correlated with problem solving performance. Factual or semantic knowledge correlated .866 with problem-solving performance, df = 18, p <.001. Procedural knowledge correlated .937 with problem-solving performance, df =18, p < .001, strategic knowledge correlated .819 with problem-solving performance, df = 13, p < .001, and rationale knowledge correlated .788 with problems-solving performance, df = 18, p < .001.

Discussion

The results of the present project demonstrate the feasibility of using CSA as a method to assess how well students have learned algebraic topics. The correlations between the assessed individual knowledge components of the INKS framework and problem-solving performance were all high and the overall correlation between the assessed overall INKS knowledge and problem solving was a near perfect .966.

The benefit of the CSA technique is its simplicity as well as its power. In the present study, CSA was implemented through a Google Forms survey. This suggests that it is readily scalable and can be implemented in educational settings with minimal disruption to existing practices and minimal training on the part of teachers. We definitely see such implementation as a logical next step of the present work.

The present research still leaves interesting questions unanswered. The present CSA technique neither measures a student's tendency to make careless mistakes nor quantifies the impact of such mistakes on the final test score. This does seem to be a surmountable problem. In Liang and Leddo (2021), software was created to probe students' knowledge after they made mistakes

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

on math problems. The software looked at the mistake that was made and then queried the student on the underlying INKS knowledge. If the student demonstrated mastery of the knowledge but still made a mistake, the software labeled the mistake as a careless one. This can be incorporated into the present framework.

Another important research direction to take is to conduct a systematic replication of the above experiment across different math subjects (e.g., pre-algebra, geometry). Of interest is not only whether the basic predictive power of the INKS framework holds but also whether different types of knowledge need to be included or different correlational strengths will emerge. Algebra 1 is highly procedural, so it is not surprising that of the four categories of INKS knowledge, procedural knowledge was the most individually predictive of overall problem solving. Geometric proofs are much more strategic in nature, so it may be the case that strategic or script-based knowledge will prove even more important for solving geometric proofs than it did for solving algebraic equations. Similarly, while algebra is highly symbolic in nature, geometry involves shapes and visual/spatial reasoning may play an important role, something not contained in the present CSA framework. CSA, and its theoretical basis INKS, may need to be expanded to incorporate this type of knowledge.

The same type of reasoning may apply to other subjects as well. Writing may incorporate extensive strategic knowledge for organizing material to be written, semantic knowledge for vocabulary and mental model/rationale knowledge for literary devices. The relative strengths of each type of knowledge may depend on the type of writing. Reading may involve even less procedural knowledge and focus more on strategic/script based knowledge for understanding the structure of text, semantic knowledge for vocabulary, and mental model/rationale knowledge to understand author's purpose. In science, semantic knowledge is be used for concepts, procedural knowledge for formulas, strategic knowledge for designing experiments, and mental model/rationale knowledge for scientific principles.

A fascinating question that arises from the present work is whether students can learn to self-assess using CSA. If this were possible, students could assess their own gaps in knowledge and then undertake corrective instruction to fill them. One challenge that may need to be overcome is that students often are not very reliable in assessing what they know and do not know. Leddo, Clark and Clark (2021) found that middle schoolers who indicated that they understood algebra content they were taught correctly answered only two-thirds of questions based on that content. Moreover, when middle schoolers indicated that they did not understand a concept, they still correctly answered three-eighths of questions based on that content. However, in the Leddo, Clark and Clark (2021) study, students were not taught how to self-assess their knowledge; they simply relied on a subjective impression of whether or not they understood the content. CSA could serve as a basis for helping students self-assess their knowledge. It may not be the case

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

that students would be able to tell if their self-assessed knowledge is accurate (although they could fact check it), but they may be able to use CSA to identify what gaps they have in their knowledge based on whether they can even answer the questions that comprise the CSA technique.

The most important research question that remains to be addressed is whether CSA, when integrated into daily classroom instruction, can boost student achievement. Here, teachers would use CSA, perhaps as part of the daily homework or in-class assignments, to assess how well students understand key concepts being taught. Any concepts that are shown to be deficient can be remediated. There is some preliminary data that suggests this may be the case. Leddo and Sak (1994) found that changes in knowledge as measured by CSA before and after instruction correlated .78 with changes in pre-test/post-test problem solving performance after instruction was given based on the initially assessed needs. However, in this study, the assessment-instruction-assessment cycle occurred just once. Future research would implement the assessment and instruction cycle on a more continuous basis.

As can be seen from the present study's results and above discussion, CSA offers a great deal of promise, both as an assessment methodology to identify what students know and how these knowledge gaps may impact performance and as part of a strategy for classroom instruction that is designed to boost student performance. Rich research opportunities are identified to address these issues.

References

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.

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

ISSN: 2455-8834

Volume:07, Issue:11 "November 2022"

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., Clark, A., and Clark, E. (2021). Self-assessment of understanding: We don't always know what we know. International Journal of Social Science and Economic Research, 6(6), 1717-1725.

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. and Sak, S. (1994). Knowledge Assessment: Diagnosing what students really know. Presented at Society for Technology and Teacher Education. .

Leddo, J., Zhang, Z. and Pokorny, R. (1998).Automated Performance Assessment Tools. Proceedings of the Interservice/Industry Training Systems and Education Conference. Arlington, VA: National Training Systems Association.

Liang, I. and Leddo, J. (2021). An intelligent tutoring system-style assessment software that diagnoses the underlying causes of students' mathematical mistakes. International Journal of Advanced Educational Research, 5(5), 26-30.

National Council of Teachers of Mathematics (2000). Principles and Standards for School Mathematics. Reston, VA: NCTM.

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

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.

