International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

AI-Generated Self-Assessment Templates Greatly Improve Math Performance in 11th Graders: A Test of the METYFoundation.org Free Self-Assessment Website

Harini Chandhu[1] and John Leddo[1,2,3]

[1]MyEdMaster, LLC, Virginia, USA
[2]METY Technology, Inc., Virginia, USA
[3]METY Foundation, Virginia, USA

DOI: 10.46609/IJSSER.2026.v11i08.046 URL: https://doi.org/10.46609/IJSSER.2026.v11i08.046
Received: 17 August 2026 / Accepted: 27 August 2026 / Published: 30 August 2026

ABSTRACT

Low student performance continues to plague educational systems around the world. The US, for example, has invested billions of dollars in educational reform over the past 50 years in the hope of dramatically raising student achievement but with little to show for this investment. Our previous research shows that Cognitive Structure Analysis-based assessment greatly improves student performance. However, these studies required us to manually create and distribute self-assessment templates to the target teachers and classrooms. In order to enable global scalability, we created a free website called metyfoundation.org that uses AI to create customized self-assessment templates for different subjects and topics, grade levels and languages. In the present study, 11th grade students studying algebra either received or did not receive the AI-generated self-assessment templates as part of the instruction supplied by their teacher.

Results showed that students who did not use the templates scored 69%, on average, on the teacher's post-test, a 2.5 percentage point increase from the teacher's pre-test, while those using the AI-generated self-assessment template scored 86%, on average, a 19.5 percentage point increase from the pre-test. This difference was significantly different. Moreover, while only half of the control students showed improvement from pre-test to post-test as a result of additional instruction, all self-assessment students improved their performance by at least 15 percentage points with the lowest scoring self-assessment group students scoring equal to the highest scoring control group students on the post-test. Results support previous self-assessment findings that a self-assessment method that takes 10 minutes to learn can dramatically improve student performance and also show that an AI-generated self-assessment template is comparably effective to a human-generated one but supports widespread scalability.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

INTRODUCTION

Throughout history, assessment has served as a measure of students' learning. Traditionally, "learning" has been defined by the number of correct answers on tests, as per classical test theory, which assumes that a student's total correct responses reflect their knowledge level (de Ayala, 2009).

Assessment methods typically fall into two broad categories: selecting correct answers from choices or constructing answers independently. Multiple-choice tests, widely used for their efficiency in grading, allow for guessing, which can inflate scores (Chaoui, 2011; Elbrink and Waits, 1970; O'Neil and Brown, 1997). Constructive response tests require students to provide their own answers, encouraging logical reasoning and offering a more accurate measure of knowledge (Herman et al., 1944; Frary, 1985). However, both methods rely on the assumption that correct answers signify learning. This assumption is problematic, as incorrect answers may point to underlying knowledge gaps, while correct answers might result from memorization or guessing, not true understanding.

Cognitive Structure Analysis (CSA) is a query-based assessment method designed to uncover the underlying knowledge concepts a student possesses, identifying the source of errors for targeted remediation (Leddo et al., 2022; Ahmad and Leddo, 2023; Zhou and Leddo, 2023; Dandemraju, Dandemraju, and Leddo, 2024). CSA is rooted in cognitive psychology research, which identifies various knowledge types, such as semantic nets (Quillian, 1966), production rules (Newell and Simon, 1972), scripts (Schank and Abelson, 1977) and mental models (de Kleer and Brown, 1981). Together, these form the INKS framework (Integrated Knowledge Structure), developed by John Leddo (Leddo et al., 1990). This framework suggests that expert knowledge is organized around scripts and principles that enable predictions and explanations.

CSA, which integrates INKS principles, has shown strong correlations with problem-solving performance: 0.966 in Algebra 1 (Leddo et al., 2022), 0.63 in scientific method problem-solving (Ahmad and Leddo, 2023), and 0.80 in precalculus (Zhou and Leddo, 2023). By assessing students' conceptual understanding, CSA enables educators to address knowledge gaps effectively, leading to significant improvements in student performance (Leddo and Ahmad, 2024; Challagulla and Leddo, 2025).

Although CSA has proven effective, the responsibility for diagnosing and remediating students' knowledge gaps lies primarily with teachers, who often manage large numbers of students. Teaching students to self-assess their knowledge could alleviate this burden. Unlike self-explanation, which involves generating explanations for learned material, self-assessment involves evaluating one's knowledge after learning.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

Cynkin and Leddo (2023) demonstrated that high school calculus students could accurately self-assess their knowledge using CSA, while Dandemraju, Dandemraju, and Leddo (2024) extended this finding to chemistry. These studies, however, addressed only the identification of knowledge gaps, not their remediation. Accurate assessment does not equate to addressing deficiencies, just as diagnosing a medical issue does not equate to treating it.

To address this issue, Ravi and Leddo (2024) conducted a study in which high school students learned an advanced topic in chemistry by watching a video. Half the students were told to rewatch the video to fill in any knowledge gaps, while the other half were taught to self-assess their knowledge using CSA and then told to rewatch the video to fill in any assessed knowledge gaps. The group that was taught to self-assess scored 15 points or 1.5 letter grades higher on a post-test than students who simply rewatched the video without self-assessment. Nehra and Leddo (2024) replicated the Ravi and Leddo study to the learning of Spanish. They found that high school students performing self-assessment plus remediation scored, on average, 25 percentage points or 2.5 letter grades higher than those re-reading the material without performing a self-assessment. Prakash and Leddo (2025a) extended the Ravi and Leddo (2024) and Nehra and Leddo (2024) findings to another subject area: reading comprehension. The results revealed a mean post-test score of 8.3 out of 12 (69.17%) for the control group and 11.2 out of 12 (93.33%) for the experimental group. Notably, individual scores further illustrated the disparity: the lowest score in the control group was 41.67%, whereas the lowest in the experimental group was 83.33%. This is the difference between an F letter grade and B letter grade. Following this, another study conducted by Prakash and Leddo (2025b) examined self-CSA's effectiveness in teaching math, specifically, the topic of Bayes' Theorem, and found a 27-point improvement. Individual scores also highlighted the disparity. The control group's lowest score was 6/20 (30%), whereas the experimental group's lowest score was 15/20 (75%). Following this, a history assessment revealed that students who utilized CSA for self-assessment and remediation significantly outperformed their peers in the control group (Prakash and Leddo, 2025c). Post-test results demonstrated that the experimental group achieved an average score of 87.5%, whereas the control group scored 65.8%, indicating a substantial difference in comprehension and retention of historical concepts.

These results on high school students were further extended by Leddo, Clark and Clark (2025) in their investigation of middle school math. Leddo, Clark and Clark found that middle school students who self-assessed using CSA and then remediated their knowledge gaps scored 18 percentage points higher on a posttest than those who relearned material without first performing a self-assessment.

Following this, Prakash and Leddo (2025d) conducted a study on middle school students' reading comprehension, specifically through an analysis of To Kill a Mockingbird, a novel that

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

explores complex themes of ethics and social structure. Students in the experimental group were trained to evaluate their own knowledge gaps and use targeted remediation strategies, while those in the control group engaged with the text without structured self-assessment. Results showed that students in the self-assessment group scored 16 points higher on a posttest than those who re-read the material without self-assessment. Building upon these results, another study examined CSA's impact on middle school students' understanding of science concepts. Students in the experimental group were taught to self-assess their understanding of key science concepts using CSA and then engage in focused review based on their assessed gaps. In contrast, students in the control group reviewed the material without guidance or structured self-assessment. Students using self-assessment scored, on average 20 percentage points or two letter grades higher on a posttest than those who did not (Prakash and Leddo, 2025e). Then, Prakash and Leddo (2025f) extended the CSA methodology to middle school history, focusing specifically on students' understanding of the causes of the American Revolution. Again, those students using self-assessment scored higher on a posttest than those who did not, this time by 29 percentage points.

Following this, Prakash and Leddo (2025g) tested whether self-assessment and remediation would work with elementary school students. This research showed that elementary school students using self-assessment and remediation for math scored an average of 83% on a posttest while those who simply reread the material scored an average of 70%. They also showed that using self-assessment and remediation raised elementary students' reading scores by an average of 20.5 percentage points (Prakash and Leddo, 2025h).

All of the above results were conducted with American students and students in K-12. Sathiyamoorthy and Leddo (2025) investigated whether self-CSA plus remediation would boost performance in college students in Scotland. Here, the testbed was college psychology. Students using self-assessment scored 15 percentages points higher than those who simply reread the material.

The studies done in the US and other countries all involved a similar format in which the experimenters administered the instructional sessions and/or prepared the instructional materials and post-test to be used. This is appropriate to establish experimental control when exploring a new topic and engaging in systematic replication. However, for the self-assessment and remediation technique to be useful in the classroom, it must also work when incorporated into a classroom's normal mode of operation without the participation of the experimenters. This was addressed by testing whether self-assessment and remediation would raise physics scores of 8th grade students in China (Chen and Leddo, 2025a). In that study, those students using self-assessment and remediation scored 23 percentage points higher than students who did not. Chen and Leddo (2025b) followed up that study with 2nd grade reading students in China and found

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

that those using self-assessment and remediation scored 15 percentage points higher on a post-test than those who did not. Challagulla, Challagulla and Leddo (2025) found that elementary school students in India who used the self-assessment method scored 22 percentage points higher on a reading post-test than those who did not.

These investigations in classrooms included having the teachers explain self-assessment to the students. While this is a reasonable activity for teachers to perform, we wanted to explore whether reducing teacher workload to the simple requirements of handing the students the self-assessment templates and checking to make sure the students filled them in (without actually analyzing the content) would still produce educational gains. Such an investigation was conducted by Gajula and Leddo (2026). In that study, an elementary school math teacher taught a unit on arithmetic operations. Prior to introducing self-assessment, the teacher gave an interim test. After that, half the students were given the self-assessment templates and told to fill them out, including what they learned after remediating their self-identified knowledge gaps, and half were not given self-assessment templates. Instruction continued for three more weeks, with self-assessment students filling out their templates once a week and the teacher checking to make sure they did. At the end of the three-week period, the teacher gave a final exam. Results showed that students in the self-assessment condition scored, on average, 18.5 percentage points higher on the final exam than did the students in the control condition. Moreover, control condition students showed no improvement, on average, over the three-week period. On the other hand, students in the self-assessment condition improved, on average, by 16.5 percentage points with all but one student raising their scores (the one student who did not show increased performance scored 90% on both the interim and final tests).

The Gajula and Leddo (2026) study was followed up by one involving high school reading (Bajwa, Leddo and Verdu, 2026). In that study, the teacher gave a reading assignment. Afterwards, half the students received the self-assessment for reading and filled it out. Both the self-assessment and no-self-assessment students had an opportunity to review the material before a test was given the next day. Those doing the self-assessment scored on average 96% on the test and improved, on average, by 29 percentage points over their previous averages in that class with each student showing score improvement. The lowest test score in the self-assessment group was 80%, which represented 40 percentage points improvement over the student's previous class average of 40%. That student did not have the largest improvement. Another student who also had a 40% average in the class scored 100% on the test given after the reading assignment. On the other hand, those who did not receive the self-assessment template, scored on average 82% on the test and showed no improvement, on average, compared to their previous averages in the class.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

The next series of field trials involved a high school in Nigeria (Ifenatuora et al., 2026). The effectiveness of self-assessment was tested in five subjects: math, chemistry, business, government and grammar. In all cases, students using the self-assessment method showed improved test performance. Average test scores for students using the self-assessment method vs. those who did not were 82% vs 64% for chemistry, 97% vs. 78% for math, 87% vs. 79% for business, 85% vs. 52% for government and 91% vs. 81% for grammar. The lowest test score for any student using the self-assessment method was 70%, which occurred in the government class- the class that had the lowest average test score for students not using the self-assessment method. Moreover, in the government class, the highest test score achieved by a student not using the self-assessment method was 68%, which meant the lowest performing student who used the self-assessment method did better than the highest performing student who did not. Additionally, a Levene's test showed that for the subjects of math, chemistry and business, the variability in scores was lower in the self-assessment group than it was in the no self-assessment group. This suggests that, in some cases, the self-assessment method not only raises student achievement but reduces the difference in performance between high and low performing students, thus helping to close the achievement gap.

Given that our experiments in Africa showed that self-assessment can reduce the variability between high and low-performers, Bajwa, Rutz and Leddo (2026) investigated whether self-assessment could reduce the achievement gap between White/Asian and Black/Hispanic students. Testing in a high school AP Psychology class, Bajwa et al. found that students using CSA-based assessment and remediation scored nearly 40 percentage points, on average, higher on the teacher's unit test than did those who did not use self-assessment. Moreover, for students not doing self-assessment, White/Asian students scored 38 percentage points, on average, than did Black/Hispanic students. However, those students using CSA-based self-assessment and remediation, Black/Hispanic students' test scores were statistically equal to those of White/Asian students and roughly two letter grades higher than the scores of White/Asian students who did not use self-assessment.

More recently, Mishra and Leddo (in press) showed that CSA-based self-assessment and remediation could even improve students' SAT scores. They found that with just a week of practice, students using self-assessment boosted their SAT Math scores by 45 points and SAT Reading by 33 points, on average, while a control group studying SAT without self-assessment showed no improvement in SAT scores.

These cumulative successes led to the creation of a free website, www.metyfoundation.org, in which people can obtain self-assessment templates, created by AI, for any subject, grade/topic, and language. The present study investigates whether a classroom using this website can see improvement in student performance.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

METHOD

Participants

Participants were 20 11th grade mathematics students who attended Sourashtra Secondary School in Madurai, India.

Materials

The educational content for the students was provided by the teacher, based on normal classroom curriculum. The topic covered was Algebra Basics and included the subtopics of one-step equations, substitution and checking answers. There was a 20-question pre-test and 10-question post-test created by the teacher.

For the self-assessment template, the teacher used the www.metyfoundation.org website. In creating the math self-assessment template, the teacher entered the following parameters: subject=math, grade=11, topic=Algebra+Basics, language=English. The self-assessment instructions that were generated are shown below:

"Self-Assessment Guide

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 basic linear equations such as $3x + 5 = 20$. 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 the concepts that describe the objects involved. For algebra basics I need to know what variables, constants, coefficients, terms, expressions, equations, and equality mean. Strategies are the general processes I would use to solve the problem. In this case the strategy is to isolate the variable by applying inverse operations in the correct order. Procedures are the specific steps that implement the strategy: first move the constant term to the other side by subtracting 5, then divide both sides by the coefficient 3. Finally, rationales are the reasons why the strategy and procedures work, such as the additive property of equality (doing the same thing to both sides keeps the equation true) and the multiplicative property of equality (dividing both sides by the same non-zero number preserves equality). With this in mind, this is how I might assess my own knowledge of solving basic linear equations. For facts, I need to know that a variable is a symbol that stands for an unknown number, a constant is a fixed number, a coefficient is the number multiplied by the variable, a term is a product of a coefficient and a variable or a constant, an expression is a combination of terms with operations, and an equation states that two expressions are equal. I'm pretty confident about most of these, but I don't remember whether the term "coefficient" also applies when the variable is raised to a power. For

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

strategies, I need to remember the overall plan of isolating the variable by undoing operations in reverse order. I know the idea of "undoing" but I'm not sure if I should always start with subtraction before division, or if the order depends on the specific equation. For procedures, I would first subtract 5 from both sides, giving $3x = 15$, then divide both sides by 3 to get $x = 5$. I'm comfortable with the subtraction step, but I'm not sure whether I need to consider the sign of the coefficient when it is negative, for example in $-3x + 5 = 20$. For rationales, I rely on the additive property of equality to justify subtracting the same number from both sides and the multiplicative property of equality to justify dividing both sides by the same non-zero number. I understand why these properties work, but I'm uncertain about how they extend to more complex expressions that involve parentheses or multiple variables.

Please use the form below to self-assess your own knowledge about a math topic. Fill in all the sections. Then when you are done, look at what you said you don't know about the topic or are unsure about and reread your lesson material to fill in your gaps."

Additionally, the form included spaces for the student to fill in the self-assessment. Each space included a description followed by a textbook for student input. The descriptions are shown below (each description had a textbox after it):

Facts: What information do you remember?

Think about definitions, vocabulary, and key information you learned.

Strategies: What methods did you learn?

Explain the method you would use, not just the answer.

Procedures: What are the steps to do it?

Can you list the steps in order? Try "First... Then... Next... Finally..." if it helps.

Reasons: Why does this work this way?

Explain the "why" behind the concept. Why is this true? Why does it matter?

Self-Assessment Summary

My Knowledge Gaps

What do I know well? What do I need to review?

What I Learned from Reviewing

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

After reviewing the material, what did you understand better?

Procedure

10 students each were assigned either to the self-assessment group or the control group (no self-assessment). All students began with the 20-question pre-test. All students received general instruction from the teacher with the addition that the self-assessment group students also performed the self-assessment. Students were then given the 10-question post-test.

RESULTS

Results showed that the mean pre-test scores for students in both control and self-assessment groups were identical at 66.5% per group. For the post-test, the mean scores were 69% for the control group and 86% for the self-assessment group. Therefore, on average, control group students improved by 2.5 percentage points and self-assessment group students improved by 19.5 percentage points. The difference in amount of improvement across groups was statistically significant, $t(18) = 9.55$, $p < .0001$.

As discussed above, students in the control group improved, on average, by 2.5 percentage points. On an individual basis, five students showed no improvement, and five students showed five percentage points of improvement. This was possible since the teacher's pre-test had 20 questions on it and the teacher's post-test had 10 questions on it. An analysis was done to compare pre-test and post-test scores to see if the 2.5 percentage points, on average, was statistically significant. It turns out it was, $t(9) = 3.00$, $p = .015$.

Students in the self-assessment group improved on average by 19.5 percentage points. At the individual student level, all students showed improvement, with the minimum improvement level being 15 percentage points or approximately 1.5 letter grades. The average improvement of 19.5 percentage points was statistically significant, $t(9) = 12.4$, $p < .0001$.

There was one other result worth noting. The control group post-test scores ranged from 60% to 80% while the self-assessment group post-test scores ranged from 80% to 100%. In other words, no experimental group student scored lower than the equivalent of a B grade while no control group student scored higher than the equivalent of a B grade, and only two control group students achieved the score of 80% on the post-test. This means that the lowest performing self-assessment group students performed at the same level as the highest performing students in the control condition.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

DISCUSSION

This program aimed to evaluate the effectiveness of self-assessment and remediation using Cognitive Structure Analysis (CSA) in helping high school students in math improve academic performance. Results showed that students not using the self-assessment method scored, on average 69% on the post-test given by their teacher, a 2.5 percentage point increase from the pre-test, while students given the self-assessment method scored, on average, 86% on the post-test, a 19.5 percentage point increase from the post-test. Both the between group levels of improvement and within group levels of improvement were statistically significant. Moreover, the lowest post-test score of students using the self-assessment method was equal to the highest score of students not using it. This result is similar to our study in Nigeria in which the lowest post-test score of high school students in the government class who used the self-assessment method was higher than the highest score of students who did not (Ifenatuora et al., 2026).

The present study, along with others cited in the Introduction of this paper, demonstrates that a simple and free self-assessment technique that takes 10 minutes for students to learn and that can be done for homework, without changing how teachers teach and test their students, dramatically raises student achievement. What the government and educational community has failed to achieve in the past 50 years can now be achieved, for free, in 10 minutes.

What makes this study noteworthy is that this is the first study in which the self-assessment template used by the students was generated using a large language model from a freely available website. This is important because such a website helps make our self-assessment method scalable globally across subjects, grades and languages. Previously, we had to create a self-assessment template (or set of templates) and email it to the teacher(s) involved for each study we conducted. This is too time intensive to create a resource that can be used globally. The present study shows that the AI-generated template can be comparably effective to the manually generated ones and, therefore, can serve as the basis for widespread use.

There is no apparent downside to implementing this self-assessment method. It produces large educational gains across the board, it takes only 10 minutes to implement, is completely free, incurs virtually no burden to teacher workload, requires no additional teacher training or change to curriculum/testing, and is politically non-controversial. Furthermore, should a teacher choose, s/he could read the students' self-assessments to see where their learning needs are and use that to inform daily teaching, and administrators could use the aggregated student self-assessment data to make data-driven decisions about curriculum design and selection of instructional materials.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

There is one area of further research that would be most interesting. Currently, after self-assessing, students perform their own remediation, which means that they have to review their instructional materials and find the relevant information to fill in their knowledge gaps. Our body of research in self-assessment shows that students can do this very well. However, in recent years, we have seen a proliferation of large language models (LLMs) and their use by students for their schoolwork. While LLMs are very powerful, they are inherently language models that provide answers to the questions they are asked. If 100 people ask the same question, it is still the same question and therefore, these people will get the same answer. However, people do not talk to each other the same way. We do not talk to children the same way we talk to adults. We do not talk to people with little knowledge of a subject area the way we talk to people who know a lot about it. LLMs do not make these distinctions.

Recently, we have been conducting experiments in which we incorporate self-assessment into LLM-based chatbots. Students perform an initial self-assessment on the topic they wish to learn about, and that information is fed to the chatbot with the instructions to tailor answers to students' questions to focus on filling self-assessed knowledge gaps. We have tested this approach on a variety of subject areas (Ganne and Leddo, 2026; Rapolu and Leddo, 2026a; Maviti and Leddo, 2025; Maviti, Leddo and Prakash, 2025; Wang and Leddo, 2025). In these tests, students learn a new subject by using either our self-assessment chatbot or a standard LLM like ChatGPT or Gemini. The instruction is followed by a post-test. We find that those using the self-assessment chatbot score, on average, the equivalent of one to four letter grades (roughly 10 to 40 percentage points) higher on the post-test than those who use a standard LLM. Rapolu and Leddo (2026b) found that while students who used self-assessment showed improvement in performance, those who used a self-assessment chatbot showed an additional 10 percentage points improvement. This suggests that AI can enhance the effectiveness of CSA-based self-assessment and remediation by assisting the remediation process. Further research could be done to see if AI can help the assessment process as well.

We have established a non-profit called METY Foundation in order to disseminate the self-assessment method free of charge to all who want to use it (www.metyfoundation.org). When fully implemented, students will be able to access self-assessments in any subject, grade and language. They will be able to print or email self-assessment reports to parents and teachers. If teachers register their classrooms, analytics on students' knowledge strengths and learning needs will be provided to support teaching. Ultimately, the self-assessment chatbot technology will be included as well to provide personalized remedial instruction.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

REFERENCES

Ahmad, M. and 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.

Bajwa, J., Leddo, J. and Verdu, M. (2026). Incorporating Student Self-assessment and Remediation into Classroom Instruction Greatly Improves Educational Performance in High School Reading Students. International Journal of Social Science and Economic Research, 11(3), 1073-1084.

Chaoui, F. (2011). The efficacy of multiple-choice tests in large classrooms. Educational Research Review, 6(3), 103-115.

Challagulla, M. and Leddo, J. (2025). Assessing and Remediating Knowledge Improves Problem-solving Performance in Middle School Math Compared to Remediating Problem-solving Steps. International Journal of Social Science and Economic Research, 10(8), 3574-3586.

Challagulla, R., Challagulla, L. and Leddo, J. (2025). Improving Reading Comprehension in Elementary School Students in India by Teaching Them to Self-Assess and Remediate Their Own Learning Needs. International Journal of Social Science and Economic Research, 10(10), 5586-5596.

Chen, Y. and Leddo, J. (2025a). Using Self-Assessment and Remediation to Raise Chinese Middle School Achievement in Physics. International Journal of Social Science and Economic Research, 10(9), 4473-4482.

Chen, Y. and Leddo, J. (2025b). Using Self-Assessment and Remediation to Raise Chinese Primary School Achievement in Reading. International Journal of Social Science and Economic Research, 10(10), 5576-5585.

Cynkin, C. and Leddo, J. (2023). Teaching students to self-assess using Cognitive Structure Analysis: Helpingr students determine what they do and do not know. International Journal of Social Science and Economic Research, 08(09), 3009-3020.

Dandemraju, A., Dandemraju, R., and Leddo, J. (2024). Teaching students to self-assess their knowledge of chemistry using cognitive structure analysis. International Journal of Social Science and Economic Research, 09(02), 541–549.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

de Ayala, R. J. (2009). The reliability and validity of standardized tests in education. Educational Measurement, 18(4), 35-50.

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, H., and Waits, B. (1970). The reliability of multiple-choice testing methods in educational assessments. Testing and Evaluation, 14(1), 54-67.

Frary, R. (1985). Constructive response tests and their impact on academic learning. Educational Measurement, 22(3), 119-132.

Gagula, A. and Leddo, J. (2026). Incorporating Student Self-assessment and Remediation into Classroom Instruction Greatly Improves Educational Performance in Elementary School Math Students. International Journal of Social Science and Economic Research, 11(2), 803-814.

Ganne, Y. and Leddo, J. (2026). Evaluating the Impact of Personalized Adaptive AI vs. Standardized ChatGPT on the Short-Term Academic Performance in Adolescents. International Journal of Social Science and Economic Research, 11(3), 1059-1072.

Herman, J. L., Klein, D. C., Heath, T. M., and 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.

Ifenatuora, G. P., Leddo, J., Sobayo, A. J., Ogude, O. F., Demian, G. C., & Emodi, U. S. (2026). Incorporating student self-assessment and remediation into classroom instruction greatly improves educational performance in senior secondary students in Nigeria. International Journal of Social Science and Economic Research, 11(4), 1622–1636.

Leddo, J. and Ahmad, M. (2024). Do You Want to Raise Student Achievement? Then, Assess and Remediate Knowledge, Not Problem-solving Performance. International Journal of Social Science and Economic Research, 9(1), 249-261.

Leddo, J., Clark, E. and Clark, A. (2025). Using self-assessment and remediation to raise middle school student achievement in math. International Journal of Social Science and Economic Research, 10(3), 1083-1092.

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.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

Leddo, J., Li, S. and 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.

Maviti, A. and Leddo, J. (2025). A Self-assessment Chatbot Greatly Outperforms Chat GPT in Teaching High School Students Calculus. *International Journal of Social Science and Economic Research*, 10(10), 5519-5531.

Maviti, A., Leddo, J. and Prakash, P. (2025). A Self-assessment Chatbot Greatly Outperforms Gemini in Teaching High School Students Calculus. *International Journal of Social Science and Economic Research*, 10(11), 6073-6085.

Mishra, G. and Leddo, J. (2026). Improving SAT Test Performance by Having Students Engage in Self-assessment and Remediation. International Journal of Social Science and Economic Research, in press.

National Assessment Governing Board. (2023). *NAEP long-term trend assessment*. https://www.nagb.gov/naep/long-term-trend.html

National Center for Education Statistics. (2023). *NAEP long-term trend assessment results: Reading and mathematics*. https://www.nationsreportcard.gov/highlights/ltt/2023/

Nehra, P. and Leddo, J. (2024). The effects of Cognitive Structure Analysis in self-assessing and remediating knowledge gaps in introductory Spanish. Journal of Educational Psychology, 45(3), 78-89.

Newell, A. and Simon, H. (1972). Human Problem Solving. Prentice-Hall.

O'Neil, H. F. and Brown, D. S. (1997). The impact of self-assessment on student performance. Journal of Educational Psychology, 12(1), 89-102.

Quillian, M. R. (1966). Semantic memory. In M. Minsky (Ed.), Semantic Information Processing (pp. 227-270). MIT Press.

Prakash, P. and Leddo, J. (2025a). Using Self-Assessment and Remediation to Raise Student Achievement in Reading Comprehension. International Journal of Social Science and Economic Research, 10(1), 277-286.

Prakash, P. and Leddo, J. (2025b). Using Self-Assessment and Remediation to Raise Student Achievement in Mathematics. International Journal of Social Science and Economic Research, 10(1), 447-456.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

Prakash, P. and Leddo, J. (2025c). Using Self-Assessment and Remediation to Raise Student Achievement in History. International Journal of Social Science and Economic Research, 10(3), 650-659.

Prakash, P. and Leddo, J. (2025d). Using Self-Assessment and Remediation to Raise Middle School Student Achievement in Reading Comprehension. International Journal of Social Science and Economic Research, 10(3), 1130-1140.

Prakash, P. and Leddo, J. (2025e). Using Self-Assessment and Remediation to Raise Middle School Student Achievement in Science. International Journal of Social Science and Economic Research, 10(4), 1471-1482.

Prakash, P. and Leddo, J. (2025f). Using Self-Assessment and Remediation to Raise Middle School Student Achievement in History. International Journal of Social Science and Economic Research, 10(6), 2429-2440.

Prakash, P. and Leddo, J. (2025g). Using Self-Assessment and Remediation to Raise Elementary School Student Achievement in Mathematics. International Journal of Social Science and Economic Research, 10(9), 4058-4068.

Prakash, P. and Leddo, J. (2025h). Using Self-Assessment and Remediation to Raise Elementary School Student Achievement in Reading. International Journal of Social Science and Economic Research, 10(10), 5532-5541.

Rapolu, S. and Leddo, J. (2026a). The Impact of a Self-Evaluation Chatbot on High School Students' Learning of Cell Communication. International Journal of Social Science and Economic Research, 11(1), 320-333.

Rapolu, S. and Leddo, J. (2026b). A Comparison of the Relative Effectiveness of Student Self-Assessment/Remediation vs. Using a Self-Assessment Chatbot on Raising Student Performance. International Journal of Social Science and Economic Research, 11(7), 3613-3629.

Ravi, D. and Leddo, J. (2024). Improving Student Performance by Having Students: Assess and Remediate Their Own Knowledge Deficiencies. International Journal of Social Science and Economic Research, 9(10), 4717-4724.

Sathiyamoorthy, S.S. and Leddo, J. (2025). Using Self-Assessment and Remediation to Raise College Student Achievement in Psychology. International Journal of Social Science and Economic Research, 10(5), 1859-1869.

ISSN: 2455-8834

Volume:11, Issue:08 "August 2026"

Schank, R. C. and Abelson, R. P. (1977). Scripts, plans, goals, and understanding. Lawrence Erlbaum.

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

