

**From F to B in 10 minutes: Self-assessment and Remediation Greatly Improves Educational Performance and Eliminates the Ethnicity Achievement Gap in AP Psychology Students**

Janat Bajwa<sup>1</sup>, Jonathan Rutz<sup>4</sup> and John Leddo<sup>1,2,3</sup>

<sup>1</sup>MyEdMaster, LLC, Virginia, USA

<sup>2</sup>METY Technology, Inc., Virginia, USA

<sup>3</sup>METY Foundation, Virginia, USA

<sup>4</sup>Gainseville High School, Virginia, USA

DOI: 10.46609/IJSSER.2026.v11i05.040 URL: <https://doi.org/10.46609/IJSSER.2026.v11i05.040>

Received: 1 May 2025 / Accepted: 26 May 2026 / Published: 30 May 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. Moreover, an achievement gap between White/Asian and Black/Hispanic students persists in which the former generally score higher on tests than the latter. Despite these interventions, education still remains a one-size-fits-all institution in which students receive the same instruction, the same textbook, the same homework and the same tests. In previous papers, we have shown that by empowering students by teaching them to self-assess their knowledge and remediate their knowledge gaps, students can improve their performance on tests by an average of 15 to 25 percentage points. The present study seeks to bolster previous findings. A high school AP Psychology class served as the testbed for this study. The teacher gave his normal lesson to a class of 16 students. Afterwards, he gave 10 students a self-assessment template and asked students to conduct a self-assessment and remediation for homework. In the next class, the teacher gave a 15-question multiple choice test. Students who were not given the self-assessment template scored, on average 42.2% on the test, while those given the self-assessment template scored, on average, a significantly higher 81.3%. Moreover, for students not receiving the self-assessment template, White/Asian students statistically significantly outperformed Black/Asian students, on average, 55% to 16.7%. On the other hand, for students who did receive the self-assessment template, the average performance of Black/Hispanic students (71.1%) was*

*statistically equal to that of White/Asian students (85.7%), thus narrowing the ethnicity achievement gap. Results support previous self-assessment findings that a self-assessment method that takes 10 minutes to learn can dramatically improve student performance.*

## **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.

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 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 8<sup>th</sup> 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 2<sup>nd</sup> grade reading students in China and found 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.

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.

The goal of the present study is to conduct an additional classroom investigation in order to further strengthen the evidence base for the efficacy of self-assessment in raising student achievement in classrooms.

## **METHOD**

### *Participants*

An AP (Advanced Placement or college level) Psychology class from Gainesville High School in Prince William County, Virginia (US) participated in the present study. There were 16 students total, 10 of which were assigned to the self-assessment condition and six were assigned to the control condition. Of the 10 students in the self-assessment condition, seven were White or Asian and three were Black or Hispanic. Of the six students in the no self-assessment condition, 4 were White or Asian and 2 were Black or Hispanic.

### *Materials*

The educational content for the students was provided by the teacher, based on his normal classroom curriculum. The topic covered was Positive Psychology, which had the components

Introduction to Positive Society, Introduction to Flow, Practicing Mindfulness, Benefits of Gratitude, Kindness and Social Connection, Healthy Practices, Resilience in Positive Psychology, and Happy Documentary. The test given was also the teacher's normal test for this unit and consisted of 15 multiple choice questions.

Because we had not constructed a self-assessment template that used a psychology example, the self-assessment template was the same one we give for math classes. We considered this to be acceptable since we had used the sample template in our prior study involving college psychology (Sathiyamoorthy and Leddo, 2025) in which students using that math self-assessment template scored significantly higher on a post-test than those who did not use the template. The template for the present study is shown below:

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 equality that 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. For strategies, I need to know 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 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.

Please use the form below to 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.

Topic you are learning:

Facts about the topic you know:

Procedures about the topic you know:

Strategies about the topic you know:

Rationales about the topic you know:

This is what you don't know about the topic or what you are unsure about

Now reread the material or the lesson and try to fill in what you don't know.

Write what you learned after rereading the material to learn what you don't know:

### ***Procedure***

Prince William County employs an A-day/B-day format in which students attend a particular class every other day. On the first day of the study, the teacher taught his normal lesson on Positive Psychology. 10 of the 16 students were given the self-assessment template and told to fill it out for homework. On the second day of the study (two days later), the teacher verified that those who were given the template did fill it out. The teacher did not evaluate or grade the quality of what students wrote in the template. The teacher then gave the class the 15-item multiple choice test.

### **RESULTS**

Upon completion of the tests, teacher scored each student's test in accordance with his normal grading practice. It should be noted that since the test employed a multiple-choice format, the grading process was objective and consistent, meaning that there was no potential for grading bias since the students' answers either matched the answer key's answers or they did not. Results showed that, on average, students who did not receive the self-assessment template scored 42.2% on the test. Students who did receive the self-assessment template scored, on average, 81.3 or nearly 40 percentage points higher. This difference was statistically significant,  $t(14) = 4.00$ ,  $p = .001$ . A Levene's test was run to examine whether the self-assessment template students' scores had less variability than those of the control group. While the variability was numerically lower, this difference was not statistically significant, possibly due to the small sample size.

There was an exploratory analysis that was done, even though the sample size was too small to draw a firm conclusion. It has been well documented that there is an achievement gap between Asian/White and Black/Hispanic students with the latter group typically performing lower in school than the former. A great deal of resources has been invested in trying to bridge this gap. If we look at how self-assessment differentially affected White/Asian vs. Black/Hispanic students, we see that in the non self-assessment condition, White/Asian students scored, on average 55% on the test, while Black/Hispanic students scored, on average, 16.7%. This difference was statistically significant,  $t(4) = 3.03$ ,  $p < .05$  and matches the standard achievement gap between White/Asian and Black/Hispanic students. On the other hand, in the self-assessment condition, White/Asian students scored, on average, 85.7% while Black/Hispanic students scored, on average, 71.1%. This difference was not statistically significant, suggesting that self-assessment can help close the achievement gap between White/Asian and Black/Hispanic students.

## **DISCUSSION**

This program aimed to evaluate the effectiveness of self-assessment and remediation using Cognitive Structure Analysis (CSA) in helping high school students in AP Psychology improve academic performance. Results showed that students not using the self-assessment method, scored, on average 42.2% on the test given by their teacher. This is equivalent to an F or failing grade by Prince William County standards. Students using the self-assessment method scored, on average, 81.3% on the test. This is equivalent to a B grade by Prince William County standards.

To put these results in perspective, in the past 50 years, the United States government and the various state and local governments have invested billions of dollars and implemented numerous educational reforms such as new curriculum frameworks, standardized testing, supplemental educational services for low income/low performing students, and providing classrooms with computers and Internet access. Despite all this, National Assessment of Educational Progress (NAEP, 2023) long-term trend data indicate that U.S. student academic performance has shown little substantial improvement over the past 40–50 years. For example, the National Assessment Governing Board (2023) reported that reading scores for 13-year-olds in 2023 were “not statistically different from scores in 1971” (NAEP Long-Term Trend Assessment). Similarly, the National Center for Education Statistics noted that long-term trend NAEP assessments have “remained relatively unchanged since 1990.”

Moreover, during this time, the ethnicity achievement gap between White/Asian and Black/Hispanic students also persisted. In the present study, White/Asian students significantly outperformed Black/Hispanic students in the non self-assessment condition, consistent with this achievement gap. However, in the self-assessment condition, the two groups’ performances were statistically equal.

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.

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.

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, 2026; 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. One area of future research is to compare educational performance of students who use self-assessment chatbot to remediate their learning needs to that of students who remediate their self-assessed learning needs without technology.

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](http://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.

## 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: Helping 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.

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.

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.

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.

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. (2026). 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.

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.

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.