# JTAR

## EDITORS

### GIL NAIZER
### APRIL SANDERS
### LAURA ISBELL
### TAMI MORTON
### SUSAN WILLIAMS

# JTAR

## Journal of Teacher Action Research
## Volume 7, Issue 3, 2021

# JTAR

## About the Journal

## Editorial Team

# HELPING STRUGGLING WRITERS THROUGH EFFECTIVE SPELLING AND WRITING STRATEGIES

**Xochitl Morales**
**Mount St. Mary's University**

**Abstract** In every classroom, some students struggle more than others in writing. When writing and spelling are used in conjunction, it requires students to develop their thoughts to produce a piece of writing. The literature review highlights how important it is for students to learn effective strategies to better support writing and spelling. Teaching students effective strategies to support writing, such as Self-Regulated Strategy Development (SRSD) and Making and Writing Words (MWW) for spelling, offer students guidance to help build their self-esteem as writers. This study was conducted in a second-grade classroom. It focused on narrative writing using both mentioned strategies. The methods used were qualitative in which students were given a pre-and post-questionnaire on how they saw and what they liked about writing. The writing intervention spanned for five weeks, including modeling, instruction, and applying both strategies. The spelling strategy was only implemented with a small group of students. The results showed significant growth in student knowledge in writing narratives, and the focus group showed improvement in their spelling application. Overall, students showed confidence as writers and an increase in their writing and spelling abilities.

**Keywords:** teacher action research, self-regulated strategy development (SRSD), making and writing words (MWW), writing, spelling, narrative writing

**Introduction**

The art of expression can use words to convey a message. In the classroom, students are taught to use pencil and paper to write down their feelings, emotions, and important thoughts. To write words, however, one needs to know how to spell. As students move beyond the phonetic stage, proper spelling becomes increasingly important. Therefore, in the classroom, teachers must teach students effective strategies to prepare students to spell and become writers. No matter how effective a strategy in these two areas might be, there are always students who will struggle. Thus, teachers must continue to look for additional strategies to help their students who need that extra assistance.

The purpose of this study was to implement strategies in both spelling and writing that would help students strengthen these areas. The research questions used to guide this intervention were: Can student writing improve by using self-regulated strategy development (SRSD)? The second question was: Does making and writing words (MWW)

strategy help improve spelling that would translate to student writing? The duration of the intervention was five weeks.

**Literature Review**

In the literature reviewed regarding research in writing, Tracy et al. (2009) described how effective writing strategies such as SRSD (Self-Regulated Strategy Development) could improve students' quality of writing. Each participating student in their study received instruction on how to use this strategy. At the end of the intervention, their writing was qualitatively better. The writing samples contained complete sentences, all story elements (characters, setting, plot, problem, solution, ending), and were longer in length. Independent student success was also measured by how well students maintained the strategy as a reference tool during writing assignments.

To analyze the process of word formation, one needs to consider the spelling strategies applied to produce each word. To facilitate word formation, students need to see how words are constructed from similar spelling root words. Rasinski and Oswald (2005) modeled the MWW (Making and Writing Words) strategy to a controlled group of elementary students. These students were able to self-identify small words found in larger, more complex structured words. Providing this type of spelling instruction helps students become familiar with spelling patterns and self-correct their writing.

*Spelling Instruction.* Spelling instruction is considered to be an essential factor that contributes to the production of writing. In the primary grades, this idea is evident due to the important skills spelling is composed of that are tied with learning to write. These skills include awareness of the number of letters and sounds in a word, common patterns for short and long vowels, and spell words with inflectional endings (Joshi et al., 2008-2009). Once a student has gained strong spelling knowledge, it often, but not always, transfers into writing. As students become more aware of the connection between spelling, forming words, and writing, they acquire orthographic knowledge to become expert writers. Graham et al. (2008) conducted a study on how well primary teachers implemented spelling instruction and made adaptations for struggling students. They initiated their study by administering a survey of 168 teachers in the United States. The sample of teachers included teachers working at both public and private schools. The community settings included urban, suburban, and rural. Based on the survey results, primary teachers taught spelling using different activities and instructional procedures to help students learn phonics, spelling rules, and strategies weekly. Fifty-seven percent of teachers reported using commercial materials to teach some aspect of spelling. Others reported using stand-alone programs and basal reading series to guide their instruction. Each approach measured achievement based on student performance. No one approach was better than the other because students were performing at or below grade level.

Fresch (2007) conducted a national survey of 296 teachers across the country in urban, suburban, and rural communities. They were concerned with showing how spelling instruction was being delivered and students' capacity to apply it to independent writing. In

this survey, many teacher respondents felt that writing words several times helped students remember, yet these words were not carried over to written work. Other findings in this survey pointed out that students often cannot spell words they know how to spell in situations other than the formal lesson. The cause for this is that students feel that once they take that weekly spelling test, those words are "buried," once the spelling test is administered, students do not realize the importance of carrying over these words into independent writing (p. 320).  Fresch (2007) further elaborated that students need to understand spelling rules and apply them to become better spellers. Thus, it will help them gain knowledge in word construction.

Graham and Santangelo (2014) conducted a meta-analytic review of studies that focused on teaching spelling to develop phonological awareness to improve spelling performance. In their analysis, each study included students in Kindergarten to 12th grade who received some form of spelling instruction. Their findings showed that students who received consistent and robust support in spelling instruction improved in phonological awareness. In the area of writing, students were able to spell more words correctly.

*Writing Instruction.* In analyzing student writers and their experiences, Lin et al. (2007) conducted a study where both developing and struggling writers in grades 2 to 8 were interviewed on how they saw themselves as writers and what a good writer does. In this study, the researchers selected one elementary and one middle school in the Pacific Northwest in an urban community. Both schools were culturally and linguistically diverse. In grades two to eight, each teacher selected four students (two developing and two struggling) from their classroom. The teachers chose the students as displaying developing or struggling writing skills. The researchers gathered their data through 20-30 minutes of one-on-one interviews with each student.

According to their findings, "novice writers are not as proficient as expert writers; novice writers are overwhelmed by transcription and working memory demands during writing" (Lin et al., p. 208). The cause for this is because novice writers are learning how to write through reading, which leads them to begin to spell words they have memorized. However, proficient writers who have had more exposure to writing have built a strong background knowledge and positive attitude towards writing. Along the same lines of creating a positive attitude for writing, proficient writers focus more on the organization and less on mechanics. Consequently, less skilled writers concentrate more on surface-level features: spelling, punctuation, and grammar (Berry, 2006).

When writing is seen either from a proficient or less proficient view, it all ties back to how well the writing instruction was delivered. A student's writing experience is based on how much emphasis the teacher places on the writing process and integration to other genres in writing (Berry, 2006). Graham et al. (2008) stated that writers who have spelling difficulties would not use words they cannot spell. When this attitude is taken from the writer, it limits the message that needs to be conveyed. It also goes back to the teacher's effectiveness in teaching spelling. Thus, teaching spelling and writing effectively has to do with the teacher's knowledge and confidence in teaching these two subject areas. Students are at

a disadvantage when there is a gap in either area, which translates into poor academic growth for struggling students and not challenging developing students to their full potential.

*Spelling and Writing Strategies.* Writing is part of the gateway for successful employment in today's society. It serves as a means to communicate effectively and transmit one's thoughts. As technology increases, writing takes on increasingly diverse forms, including texting, messaging, and blogging. For someone to use these new ways of communication, good writing skills and strategies are needed.  Tracy et al. (2009) conducted a study where 127 third-grade students from a mid-western elementary school received writing instruction in their general education classroom. The 127 students were from six classrooms, where three classrooms were selected to receive SRSD strategy instruction, and the rest received traditional writing instruction. Both groups were administered the Test of Written Language (TOWL3) before writing intervention using the SRSD strategy. Previous state standardized scores for each student were analyzed to see any discrepancies in student performance. The SRSD model consists of the following: 1.) develop background knowledge, 2.) discuss the strategy, 3.) model strategy, 4.) memorize strategy, 5.) support/ scaffold the strategy. 6.) encourage independent use (p. 194). Participating students who received SRSD showed improvement in writing performance, demonstrating that teaching students strategies and highly scaffolded procedures can lead to successful results.

The techniques used by teachers implementing the SRSD strategy required students to memorize specific acronyms to help them apply the writing strategy: one of these is the POW strategy (Pick my idea, Organize my notes, Write and say more) (Tracy et al., 2009). This mnemonic device allows students, both developing and poor writers, to think about an idea first before writing. A graphic organizer is used to organize their notes and ideas and develop clear thoughts to write a cohesive essay. The last focus area is to encourage students to write more and expand their ideas once written down on paper.

Once students feel confident and show POW mastery, they can apply genre-specific strategies to their writing, such as WWW +2H, which stands for 1). Who are the main characters? 2). When does it take place? 3). What happens next? 4). How do the main characters feel? 5). How does it end? This strategy is designed to be used when writing a story. It allows teachers to instruct students in identifying story elements in their writing, such as the sequence of events, inferring character feelings, and identifying character traits. Another strategy that bolsters writing is focused on teaching students how to spell by making words. Rasinski and Oswald (2005) collaborated on a two-year project conducted in a second-grade classroom. The participating teacher (Ruth Oswald) introduced this strategy in her classroom and another partner (a second-grade teacher) at the same school site. Oswald implemented the MWW (Making and Writing Words) strategy. Only nine students were selected as being high achieving (3), average (3), and struggling (3). The other participating teacher used the district's adopted basal program and selected nine students using the same criteria.

With this spelling strategy, students are given vowel letters and consonant letters to write on a graphic organizer. The teacher then gives students clues using the letters written to make new words. Through the lesson, students are encouraged to see how each word is formed and what letters are used in each word. Students then notice how each word contains similar letters. An example given by the teacher, Oswald, is when students formed the word "hate" and compared it to heat. Some noticed that /ea/ does not always have a long sound "head."

Another example was a spelling error one student made and independently corrected. The word was "becase," which the student visually identified as not looking right and fixed it. With this strategy, students were able to see little words in huge words, for example, "sea" add "m," and it makes "seam" (Rasinski & Oswald, 2005). Having students take part in their learning of language through scaffolded instruction helps facilitate understanding of how language works—resulting in spelling knowledge to be transferred into other curricular areas where writing is required.

*Assessing Student Spelling and Writing.* Once a child has mastered spelling and writing strategies, a teacher's job is to assess students' errors. To determine a child's spelling errors, sufficient samples need to be collected (Apel & Masterson, 2001). In a case study conducted by Apel and Masterson (2001), a 13-year-old student demonstrated low self-esteem, always complained about school, and felt embarrassed about her performance. She avoided writing due to poor spelling skills. She lacked phonemic awareness and orthographic skills, according to her tests. The lack of phonemic awareness for this student's writing was evident in errors of omission (example: "sop" for "stop). To address such orthographic skills, instruction should focus on teaching spelling strategies and finding words with similar patterns to help create orthographic images of words. This case study showed that the implementation of specific spelling intervention proved to be successful in targeting deficient skills in spelling. The student in this study was able to gain phonemic and decoding skills to aid in spelling and decoding unfamiliar words. The success also involved the modeling and scaffolding of new strategies that focused on the student's deficient skills.

**Methodology**

The purpose of this study was to see how small-group spelling instruction and instruction in self-regulated writing strategies can help struggling writers improve. Through effective instruction of explicit strategies in writing and spelling, struggling spellers gained more confidence in their writing, and all students improved in the overall quality of their writing. The two strategies discussed in this literature review were implemented in a second-grade classroom. The two strategies are SRSD (Self-Regulated Strategy Development) and MWW (Making Words Work).

The rationale for selecting these two strategies was to address concerns in the quality of students' writing and spelling. For the spelling strategy, not all students received the intervention. The intervention was for students whose spelling errors were severe and

impeded the deciphering of words. The SRSD strategy was implemented to encourage students to improve their writing by including enough details to elaborate on their ideas. The students in this study needed a strategy to boost their confidence and not shut down when asked to produce a writing piece.

*Participants.* This study's participants were a convenience sample of second-grade students, which included 12 girls and 7 boys.  The socio-economic status of all students is lower-middle working class. All students were included in the writing strategy portion of the study. For the study's spelling strategy portion, a focus group of four students received small group instruction. These four students struggled with both spelling and writing. These participants were specifically selected based on weekly spelling test scores in which they routinely score 50% or lower. The focus group's composition is as follows: one student has ADHD and is taking medication, one has shown signs of dyslexia (though not diagnosed), and two have a learning disability in processing the information as stated in their IEP.

*Intervention.* There were two components to the intervention: a whole class writing strategy intervention and a supplemental spelling strategy focused only on struggling spellers. The writing consisted of a personal narrative, including a beginning, middle, and end. Students were expected to add details and write these in sequential order. The focus group received additional spelling instruction.

The first step in identifying with which spelling patterns the focus group had difficulties, a word inventory was used to identify these patterns. Next, the MWW strategy was introduced to encourage students to think of how words are made from other words. During each week of the intervention, students were given specific letter tiles to sort out words with similar orthographic patterns and used appendix D. Students sorted letter tiles that contained the letters for the weekly spelling words into vowels and consonants. To illustrate this idea, let say students had ten spelling words, and only 15 letters can be used to form all words, then students could only use the letters given. Next, the teacher provided students with clues using the weekly spelling list. Then, students using the given letter tiles listened to each clue and placed the corresponding letter tiles in the order given to form the word. The clues helped students connect sounds to letters and develop phonemic awareness. The exact process was repeated for each spelling word and placed next to each numbered box in appendix D.

For the intervention's writing, all students were taught story structure using the WWW+2H mnemonic strategy to identify all the elements. A reading selection from the reading program was selected. The teacher taught a short lesson on story structure using appendix F and modeled how to fill in each section.

*Week one*. During the first week, the pre-assessment writing prompt (see Appendix A) was administered to all students. The assessment asked students to write about something special they did with a friend. Students were encouraged to include as many story elements as they remembered. To self-report how many story elements they had included in their story, students were given a story rocket graphic organizer (see Appendix B) and

were asked to fill this out. The organizer was explained previously and used to help them monitor their story writing. The pre-assessment story was graded by using a 4-point rubric (see Appendix C). All students were given a ten-question survey (see Appendix D) to measure their attitudes towards writing. The focus group was also assigned a five-question survey (see Appendix I) to measure their spelling attitudes.

The school's adopted language arts curriculum was used to help students understand story structure using the reading unit's text selections. As part of the phonics section, the focus was on words with the suffix -er. A spelling lesson was implemented to introduce and decode the words as outlined in the reading program. Following the lesson, students were given a practice spelling test with words using the -er suffix. This test aimed to see what words students could spell already and which ones they needed to study. After this, the focus students (low-performing spellers) were re-taught the -er spelling pattern in a small group setting using the MWW (Making Words Work) strategy. Students were given a graphic organizer (see Appendix D) to complete this activity.

Next, a teacher-led discussion was used to introduce the following two strategies as part of SRSD: POW (Pick my idea, Organize my notes, and Write) and WWW+What2+How2 (Who is the main character, When does the story take place. Where does the story take place, What do the characters do, What happens then, How does the story end, and How do the characters feel) (see Appendix E). These mnemonic devices were placed on chart paper for visual purposes and explained to students.

To provide students practice with the strategy, a graphic organizer (see Appendix F) was used along with a copy of a teacher-selected story. Students read along silently while the teacher read the story out loud. Students were then asked to identify Who, When, and Where student responses were recorded under the graphic organizer's appropriate space. The teacher modeled writing phrases instead of full sentences to help students get the idea down on paper and later translate these ideas (phrases) into sentences. This routine continued for the What2 and How2 parts. Students were given another short story to help them transfer the strategy. This time students were paired with a partner and given a new graphic organizer to identify all seven parts of a story.

*Week two*. In the second week, the same strategy for spelling MWW was implemented just with the focus students. The lesson's focus was on using words with contractions. Students were provided with the MWW graphic organizer to complete (see Appendix D). To see if all students remembered the strategies introduced, the teacher reviewed the charts. Prompting was used when needed. The students were reminded of these terms each day to make sure they had them memorized.

*Week three.* In the third week of the intervention, all students' spelling focus was writing and reading words that end with -y. The focus group practiced the spelling pattern using the MWW strategy and graphic organizer (see Appendix D). A copy of the POW and WWW+What2+How2 graphic organizer was displayed on a document camera for students. The teacher introduced the day's lesson by telling students, "Remember that the first letter

in POW is P-pick my idea. Today we are going to practice how to think of a good story idea and good story parts. To do this, we have to be creative." The second letter in POW was reintroduced, which is O-ORGANIZE. To organize the story, students reviewed WWW+What2+How. The final letter in POW, which is W-Write, was modeled. As the story was written, students were told to use "million-dollar words" in their writing. These descriptive words help students describe an object, place, or character. To conclude the lesson as a whole class, the story elements used were graphed by completing the story rocket- appendix B.

*Week four.* Week four was the final week for the spelling strategy MWW to be implemented with the focus group. The focus was on words ending with -er and -ing. The MWW organizer was used. After presenting all the parts of the SRSD writing model to students, at this point were given practice in applying the strategy. First, the teacher reviewed and reminded students of the strategy and verbally stated the expectations as they wrote their stories. Students were given the choice of writing a story on any topic for this assignment. After they were done, they assessed their writing piece and graphed the story elements they used by completing the story rocket graphic organizer (Appendix B). As noted previously, this organizer allowed students to count the number of story elements included in their own stories and color in that number on a graph. The goal was to increase the number of boxes in each graph to fill all boxes. Students were encouraged to count the number of words they used, write the number on the top right-hand corner of the paper (Appendix G), and color the corresponding number. Consequently, the stories students wrote had more words and details.

*Week five.* In the last week of the study, all students were assessed on how well they used the POW and WWW+ What2+How2 strategies when writing. They were given a formal writing prompt (see Appendix H) and were required to include all seven-story elements. They were given another story rocket graph (see Appendix B) to monitor their writing independently. All students' writing was graded on a 4-point scale rubric (see Appendix C) to measure how well they mastered the strategy and story writing.

**Results and Discussion**

*Writing Assessment.* Table 1 shows the results obtained from the writing assessment given to students both before and after the intervention period. This assessment consisted of a writing prompt where students were given a topic. In this case, they had to write about spending time with a friend (pre-assessment) and about a time they were brave (post-assessment). During the administration of each writing prompt, students were presented with stories related to the topic (i.e., friendship or bravery) to help build students' background knowledge. Each student's writing sample was graded using a four-point rubric (see Appendix C), and average scores for each category on the rubric were calculated for both the entire class (see Table 4.1 below) and focus students (see Table 4.2 below).

*Table 1: Pre-and Post-Intervention Writing Prompt Average Scores Per Rubric Category: Entire Class (N=19)*

| Rubric Category | Baseline (out of 4 pts. Total) | Intervention | Change |
|---|---|---|---|
| Writing Process | 2.4 | 3.4 | +1.0 |
| Focus on Topic | 2.6 | 3.4 | +0.8 |
| Organization | 2.4 | 3.6 | +1.2 |
| Punctuation | 2.1 | 3.4 | +1.3 |
| Spelling | 2.0 | 3.5 | +1.5 |

Students scored an average between 2.0 and 2.6 points out of 4.0 points on each element of the rubric on the pre-assessment writing prompt. The lowest average score (2.0) was in the area of spelling, while the highest average score (2.6) was in the area of staying focused on topic. After the intervention, students' average scores changed in all categories, as shown above, with a total average of 3.5. On the post-intervention writing prompt, the highest average score (3.6) was in the organization, while the lowest average scores (3.4) were in the writing process, focus on the topic, and punctuation. As indicated in the final column, each average score changed for the better, with scores in most areas improving by at least 1.0. Average scores for the focus group students on each rubric element were also calculated (see Table 2 below).

*Table 2: Pre-and Post-Intervention Writing Prompt Average Scores Per Rubric Category: Focus Group (N=4)*

| Rubric Category | Baseline (out of 4 pts. Total) | Intervention (out of 4 pts. Total) | Change |
|---|---|---|---|
| Writing Process | 1.8 | 3.3 | +1.5 |
| Focus on Topic | 2.3 | 3.3 | +1.0 |
| Organization | 1.8 | 3.5 | +1.7 |
| Punctuation | 1.5 | 3.3 | +1.8 |
| Spelling | 1.0 | 3.5 | +2.5 |

Table 2 above shows average scores from the focus group students on each element of the rubric. Before the intervention, the lowest average score of 1.0 was in spelling; after the intervention, the average score in this area went up to 3.5. There was growth in all other areas of the rubric as well.

*Table 3: Pre- and Post-Intervention Writing Prompt: Average Number of Story Elements Used*

|  | Baseline | Intervention | Change |
|---|---|---|---|
| **Entire Class** | 4 | 7 | +3 |
| **Focus Group** | 3 | 7 | +4 |

Table 3 above shows the average number of story elements students included in their writing before and after the intervention. On the baseline assessment, the entire class averaged the use of four out of the seven elements. After the intervention, they were able to include all seven. As for the focus group, during the baseline assessment, they had three out of the seven elements, which was one less than the entire class's average. After the intervention, their stories also included all seven elements.

*Spelling Assessment.* In addition to implementing the writing intervention, a focus group of struggling students in spelling received additional support. Table 4 below shows their average pre-intervention and post-intervention spelling test scores using the MWW strategy.

*Table 4: Focus Group Pre- and Post- Intervention Average Spelling Test Scores: Focus Group Students (N=4)*

| Student | Baseline (out of 100 pts) | Intervention (out of 100 pts) | Change in pts. |
|---|---|---|---|
| **Josh** | 86.5 | 89.5 | +3 |
| **Sal** | 62.2 | 78.5 | +16.3 |
| **Chris** | 67.7 | 65.8 | -1.9 |
| **Abby** | 42.7 | 44.8 | +2.1 |

Table 4 above shows improvement for three of the four students in their pre-and post-intervention spelling test scores. Sal improved the most by 16.3 points. Chris did not improve; instead, his score decreased by 1.9 points.

*Survey Results.* Table 5 below shows results from a survey all students were given both before and after the intervention. The survey's purpose was to see how students felt towards writing and what part of writing specifically they liked or did not like. This survey consisted of ten questions, asking students to check "yes," "sometimes," or "no" for each one.

*Table 5: Writing Survey Results: Pre-Intervention and Post-Intervention (N=19)*

| | Yes | Pre-Intervention Sometimes | No | Yes | Post-Intervention Sometimes | No |
|---|---|---|---|---|---|---|
| 1. I like writing stories. | 9 | 9 | 1 | 17 | 2 | 0 |
| 2. Writing is boring. | 1 | 6 | 12 | 0 | 3 | 16 |
| 3. I like writing at school. | 16 | 3 | 0 | 17 | 2 | 0 |
| 4. I have trouble thing about what to write. | 4 | 13 | 2 | 4 | 7 | 0 |
| 5. Writing is fun. | 13 | 5 | 2 | 4 | 7 | 8 |
| 6. I think I'm a good writer. | 12 | 6 | 1 | 16 | 3 | 0 |
| 7. I like to share my writing with others. | 12 | 7 | 0 | 14 | 5 | 0 |
| 8. I like to write about things I have learned. | 15 | 4 | 0 | 18 | 1 | 0 |
| 9. It helps me to have someone read over my writing. | 16 | 0 | 3 | 19 | 0 | 0 |
| 10. I like to think about ideas before I write. | 16 | 3 | 0 | 17 | 2 | 0 |

Overall, the survey results show that students like to write and think writing is fun. On the pre-intervention survey, responses for questions 1 and 5 fell heavily under - "yes" and "sometimes." After the intervention, the scores in these same two questions fell under "yes" the most. The two questions whose responses did not change as much were questions 6 and 7. These questions asked students if they thought they were a good writer and liked to share their writing. The table shows twelve responded "yes" for both questions (pre-intervention) and answered sixteen "yes" (for question 6) and fourteen for (question 7) post-intervention. The remaining questions in the survey showed a positive difference in students' attitudes after the intervention.

An example is that students found it easier to write their story using the strategy instead of not knowing what to write (question 4). Another example is that students enjoyed writing stories. As shown in the table above (question 1), their "yes" responses nearly doubled after the intervention. Students, in general, felt better and more confident about their writing. In addition to completing the writing survey, the focus group of four "struggling speller" students received another five-question survey on their attitude towards spelling (see Table 6 below), with each question asking students to circle "yes," "sometimes," or "no."

*Table 6: Focus Group Spelling Survey Results: Pre-Intervention and Post- Intervention (N=4)*

|  | Yes | Pre-Intervention No | Sort of | Yes | Post-Intervention No | Sort of |
|---|---|---|---|---|---|---|
| **1. Are you a good speller?** | 3 | 0 | 1 | 4 | 0 | 0 |
| **2. What do you do when you don't know how to spell? Do you sound it out?** | 1 | 0 | 3 | 3 | 1 | 0 |
| **3. Do you like spelling?** | 4 | 0 | 0 | 3 | 0 | 1 |
| **4. Do you study for spelling tests?** | 3 | 0 | 1 | 3 | 0 | 1 |
| **5. Do you like writing words?** | 4 | 0 | 0 | 4 | 0 | 0 |

The results from this survey show that the focus students' confidence towards their spelling (as indicated in question one) increased after the intervention, given that all four students at that point said, "Yes, I am a good speller." On Question 2 before the intervention, when students were asked what they did when they couldn't spell a word, most (3 out of 4) of them answered that they sort of sounded it out. After introducing the intervention, this changed to mostly "yes" (3 out of 4 students). Responses to the last three questions in the survey remained the same before and after the intervention. As for the third question, when asked if they like to spell during the pre-intervention, all students answered "yes" compared to three "yes" and one "sort of" for the post-intervention.

**Conclusion**

The focus of this study was to implement effective writing and spelling strategies. The collected data from both writing assessments and surveys proved that explicit instruction in research-based strategies helped improve student writing and spelling abilities. All participating students' attitudes towards writing changed due to their knowledge and proficiency in writing stories. This change was evidenced by student confidence in knowing how to write a narrative. Students found the mnemonic device helpful and easy to remember.

Most importantly, students were capable of applying SRSD.  This study's results can help future and current teachers implement effective writing and or spelling strategies to guarantee success. For educators who might not feel comfortable enough to teach writing, this strategy provides explicit instructions on how each component works and should be taught.  Students of all learning levels can benefit and strengthen their writing if they have a strategy they can use.

**About the Author**

**Xochitl Morales** holds a bachelor's degree in liberal studies from Cal State University Northridge, two master's degrees, one in child and adolescent literacy from Loyola Marymount University, and the other in education from Mount St. Mary's University. She is currently working on her doctoral degree from Chapman University with an emphasis on cultural and curricular studies. She is presently a full-time dual-language Spanish Kindergarten teacher and part-time college instructor. A strong advocate in creating a curriculum that includes all learners and incorporates problem-solving and critical thinking skills. A believer in equitable opportunities for all students in order to achieve success. Her professional interests focus on maintaining one's identity through first language maintenance, encouraging students to build an interest in literacy, and preparing teacher candidate students for fieldwork at the elementary level. She has been an educator for the past 15 years in grades K-5 and at the college level. Email: xmorales@chapman.edu

## References

Apel, K. & Masterson, J., J. (2001). Theory-guided spelling assessment and intervention:  A case study. *Language, Speech, and Hearing Services in Schools*, 32, 182-195.

Berry, R. A. (2006). Beyond strategies: Teacher beliefs and writing instruction in two primary inclusion classrooms. *Journal of Learning Disabilities*, 39(1), 11-24.

Fresch, M. J. (2007). Teachers' concerns about spelling instruction: A national survey. *Reading Psychology*, 28, 301-330.

Graham. S., Morphy, P.. Harris, K.R., Chorzempa. B. F., Saddler, B., Moran, S. & Mason, L. (2008). Teaching spelling in the primary grades: A national survey of instructional practices and adaptations. *American Educational Research Journal,* 45 (3), 796.

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703-1743.

Joshi, R., Treiman, R., Carreker, S., & Moats, L. (2008-2009, Winter). How words cast their spell: Spelling is an integral part of learning the language, not a matter of memorization. *American Educator.* http://www.aft.org/sites/default/files/periodicals/joshi.pdf

Lin, C. S., Monroe, B. W., & Troia, G. A. (2007). Development of writing knowledge in grades 2-8: A comparison of typically developing writing and their struggling peers. *Reading and Writing Quarterly*, *23*, 207-230.

Rasinski. T., & Oswald, R. (2005). Making and writing words: Constructivist word learning in a second-grade classroom. *Reading and Writing Quarterly*, 21 (2), 151-163.

Tracy, B., Reid, R., & Graham, S. (2009). Teaching young students strategies for planning and drafting stories: The impact of self-regulated strategy development. *The Journal of Educational Research*, 102 (5), 323-331.

**Appendix A:  Pre-Assessment**

Directions: Think about a time you did something special with a friend.

Write a story that tells what you did with your friend. Include details such as: what happened, when did it happen and where it happened.

Checklist:

*Write in complete sentences
*Add details to your writing
*Include beginning, middle, and end
*Start a sentence with a capital letter and end with a period

**Appendix B:  Story Rockets**
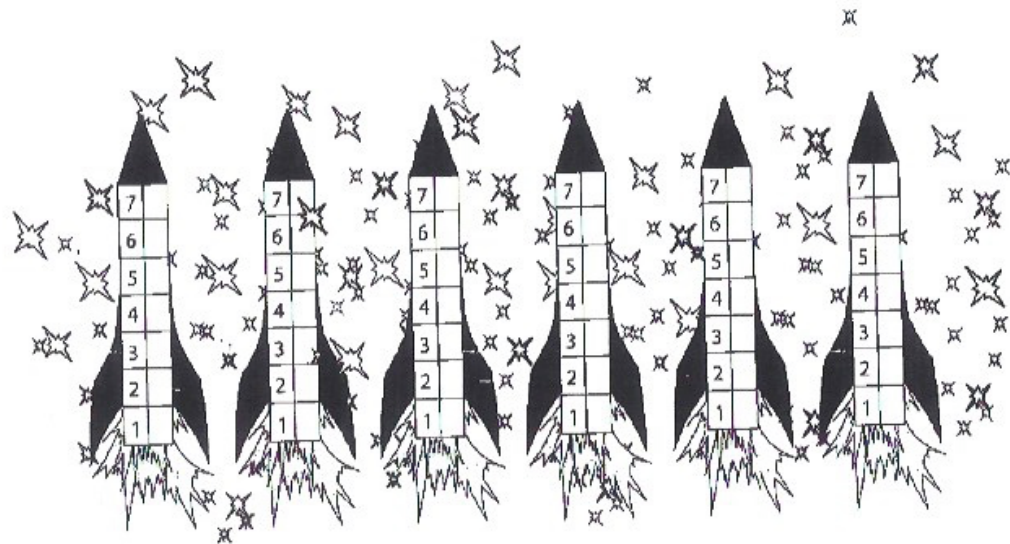


Story Rockets

Appendix B

47

**Appendix C:  Story Writing: Rubric for Pre- and Post-Assessment**

48

## Story Writing : Rubric for Pre- and Post-Assessment

Teacher Name: **Ms. Morales**

Student Name: _____

| CATEGORY | 4 | 3 | 2 | 1 |
|---|---|---|---|---|
| Writing Process | Student devotes a lot of time and effort to the writing process (prewriting, drafting, reviewing, and editing). Works hard to make the story wonderful. | Student devotes sufficient time and effort to the writing process (prewriting, drafting, reviewing, and editing). Works and gets the job done. | Student devotes some time and effort to the writing process but was not very thorough. Does enough to get by. | Student devotes little time and effort to the writing process. Doesn't seem to care. |
| Focus on Assigned Topic | The entire story is related to the assigned topic and allows the reader to understand much more about the topic. | Most of the story is related to the assigned topic. The story wanders off at one point, but the reader can still learn something about the topic. | Some of the story is related to the assigned topic, but a reader does not learn much about the topic. | No attempt has been made to relate the story to the assigned topic. |
| Organization | The story is very well organized. One idea or scene follows another in a logical sequence with clear transitions. | The story is pretty well organized. One idea or scene may seem out of place. Clear transitions are used. | The story is a little hard to follow. The transitions are sometimes not clear. | Ideas and scenes seem to be randomly arranged. |
| Punctuation | There are no punctuation errors in the final draft. | There is 1 punctuation error in final draft. | There are 2-3 punctuation errors in the final draft | The final draft has more than 3 punctuation errors. |
| Spelling | There are no spelling errors in the final draft. Character and place names that the author invented are spelled consistently throughout. | There is one spelling error in the final draft. | There are 2-3 spelling errors in the final draft. | The final draft has more than 3 spelling errors. |

## Appendix D:  Making and Writing Words

Making and Writing Words – Appendix D

| Vowels | Consonants |
|---|---|
|  |  |
| 1 | 5 |
| 2 | 6 |
| 3 | 7 |
| 4 | 8 |

Transfer

| T-1 | T-2 |
|---|---|
| T-3 | T-4 |

Fr: Rasinski, T. (1999). Making and writing words. *Reading Online*. Available at http://www.readingonline.org/articles/rasinski/. Permission to photocopy for educational use is granted. See. Making and Writing Words. Rasinski and Heym.    http://www.shelleducation.com/rasinski.php

**Appendix E:  POW Strategy**

# POW

**P**ick my Idea
**O**rganize my Notes
**W**rite and Say More
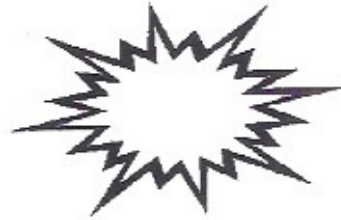
## W-W-W  What=2  How=2

**W**ho is the main character?

**W**hen does the story take place?

**W**here does the story take place?

**W**hat does the main character do or want to do; what other characters do?

**W**hat happens then?  What happens with other characters?

**H**ow does the story end?

**H**ow does the main character feel; how do other characters feel?

**Appendix F:  Story Elements Organizer**

## POW + W-W-W

WHATₐ2
HOWₐ3

| WHO | WHEN | WHERE |
|-----|------|-------|
|     |      |       |
|     |      |       |
|     |      |       |
|     |      |       |
|     |      |       |

| WHAT | WHAT | HOW | HOW |
|------|------|-----|-----|
|      |      |     |     |
|      |      |     |     |
|      |      |     |     |
|      |      |     |     |
|      |      |     |     |
|      |      |     |     |
|      |      |     |     |

**Appendix G:  Number of Words**
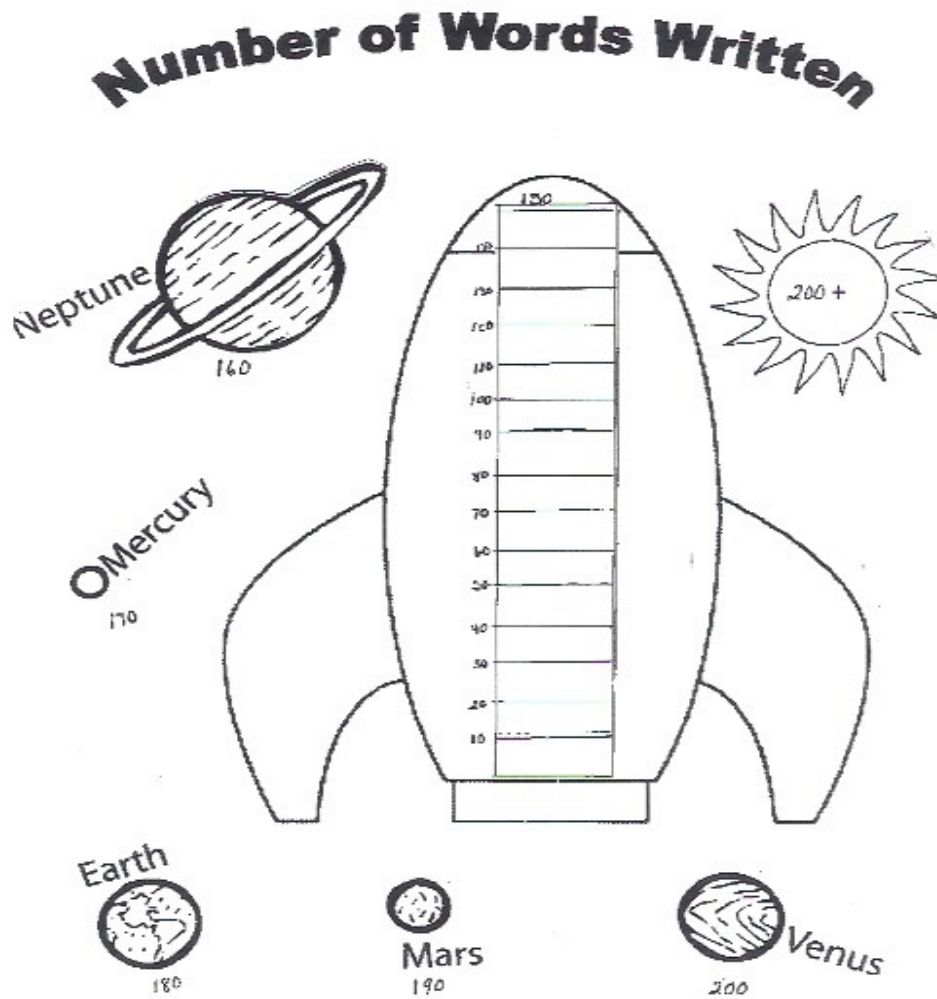
**Appendix H:  Post-Assessment**

Directions: Think about a time you helped someone, or someone helped you.

Write a story that tells what you did and includes details such as: what happened, when did it happen, and where it happened.

Checklist:

*Write in complete sentences
*Add details to your writing
*Include beginning, middle, and end
*Start a sentence with a capital letter and end with a period

## Appendix I: Spelling Interest Survey

Name: _____

    Date:_____

Spelling Interest Survey

Circle the best answer for each.

1. Are you a good speller?

Yes        No        Sort of

2. What do you do when you don't know how to spell do you sound it out?

Yes        No        Sometimes

3. Do you like spelling?

Yes        No        Sort of

4. Do you study for spelling tests?

Yes        No        Sometimes

5. Do you like writing words?

Yes        No        Sort of

# IMPROVING PRONUNCIATION SKILLS OF GRADE 6 PUPILS THROUGH ORAL DRILLS

**Antonio Jr. T. Enerio**
**Northwestern Mindanao State College**
**of Science and Technology**

**Abstract** This study explored the use of oral drills to improve the pronunciation skills of Grade 6 pupils. Word lists were used for oral drills and a semi-structured interview to capture their experiences. The drills focused on the least pronounced consonant and vowel sounds, which lasted for ten weeks. The data revealed that the learner's pronunciation skills were Satisfactory in the pretest and Very Satisfactory in the post-test. The t-test result suggested the drills positively influenced pronunciation illustrated with increased scores on the post-test. Data suggests that learner pronunciation skills significantly improved with the implementation of oral drill strategies.

## Introduction

When the researcher taught in the Department of Education, learners exhibited difficulty with proper English pronunciation. They were unable to express ideas, both written and orally. Communication involves three steps: thought, encoding, and decoding; many pupils remained at the initial step. They had the intellect but were unable to encode and decode it. Teaching them how to speak English fluently has been a complicated process, considering the three areas of knowledge: mechanics, functions, and social and cultural rules and norms. Thus, the researcher focused on mechanics, which involves pronunciation, grammar, and vocabulary. However, mechanics seems complicated; knowing the students' pre-existing knowledge-base, developing simultaneously all three may require much effort and more extended time; hence, this research focused on pronunciation.

This study assessed the pronunciation skills and introduced oral drills to the Grade 6-Earth pupils for SY: 2018-2019 of Aloran Central School, Aloran District in the Division of Misamis Occidental, Northern Mindanao, Philippines. The specific objectives were to:
1. Identify the level of the learner's pronunciation skills in the pretest;
2. Identify the level of the learner's pronunciation skills in the post-test; and
3. Determine the significant difference between the level of the learner's pronunciation skills in the pretest and post-test.

## Literature Review

Pronunciation is the manner of correctly speaking a word. Some English words are not spoken the way they are written, and more than one combination of letters represented some sounds (Nordquist, 2016). The first step in understanding the proper pronunciation of English is to concentrate on individual sounds, called phonemes. Every word is made up of many phonemes (Beare, 2017).

At the beginning level, English learners should focus on the basics of pronunciation. Rote learning is best for this level. Teaching the International Phonetic Alphabet (IPA) is a challenge at this stage, as learners are already overwhelmed by the complexities of learning a language (Nordquist, 2017). Learning another alphabet for pronunciation is challenging for most beginning-level English learners.

In improving English pronunciation, it is essential to understand some terms and concepts, according to Beare (2017): (1) Phoneme. It is a unit of sound. These are expressed as phonetic symbols in the International Phonetic Alphabet (IPA). (2) Letter. The English alphabet has twenty-six letters. Depending on letter combinations, letter pronunciations differ. (3) Consonants. These are the sounds that interrupt vowel sounds. These are combined with vowels to form a syllable. (4) Vowels. These are open sounds caused by the vibration of vocal sounds but without obstruction. (5) Consonants interrupt vowels to form syllables. All vowels are produced using the vocal cord, hence considered as voiced. (6) Voiced. A voiced consonant is formed using the vocal cords. A better way to tell if a consonant is voiced is to touch one's fingers to the throat. If the consonant is voiced, one will feel a vibration. A voiceless consonant is produced without the help of the vocal cords. Place one's fingers on the throat when speaking a voiceless consonant, and one will only feel a rush of air through the throat. (7) Syllable. It is formed when consonant and vowel sounds are combined. Words can have one or several syllables. To test the number of syllables, put a hand under the chin and speak the word. Each time the jaw moves, indicate another syllable.

When one learns a foreign language in early childhood, there is a likelihood that one learns to speak it fluently and often without any noticeable local accent. As one grows older, the chance becomes so slim and, to a large extent, impossible to acquire a native-like accent (Moeller & Catalano, 2015). At first glance, it seems that learning a second language pronunciation should be easy, a simple matter of imitating native speakers' pronunciation. However, the reality is quite different. A learner's native language affects second language pronunciation learning when the native language's sound system is different from the target language's sound system. A learner's pronunciation errors may be affected if a particular sound does not exist in the native language. The learners cannot form it; they try to substitute the nearest equivalence they know. It may also be affected if a sound exists in the native language, but not as a separate phoneme, and learners do not perceive it as a distinct sound; thus, they fail to pronounce it correctly (Long, 2011).

Adults are unlikely to attain a native-like accent, while both empirical and anecdotal findings attest that children can acquire a target-like proficiency in pronunciation. The cut-off age

ending this period in pronunciation acquisition was formerly believed to be around 13. However, recent studies substantiate that even after the age of 6, the learner's accent is discernible to be non-native, albeit with slight discrepancy (Smith & Candlin, 2014). The impact of age on pronunciation learning has suggested that it is brought about by progressive decline rather than a substantial drop-off after puberty. The positive thing is that these neurological variations between adults and children tend to be the product of a transition rather than a decline. In the way that sounds are processed in a foreign language, training will help adults develop their ability to differentiate against new sounds and create new phonetic boundaries (Long, 2011).

Understanding and teaching pronunciation has been controversial in second language acquisition for many years. Teachers drilled learners until they had the appropriate accent. To be understood meant capturing one of the established accents.  There is now an emphasis on the comprehensibility of what they say. According to Gilakjani (2012), one of the critical requirements for language proficiency is to secure understandable pronunciation for language learners. Teachers must act as pronunciation coaches, and learners must be proactive learners taking the initiative to learn.

One of the goals of teaching pronunciation in any course is intelligible pronunciation, not perfect pronunciation. Being able to attain perfect pronunciation should no longer be the objective. However, more realistic goals should be reasonable, applicable, and suitable for the learner's communication needs. The learner needs to develop: functional intelligibility, which is the ability to make oneself easily understood; functional communicability, the ability to meet the communication needs one has to face; improved self-confidence; and the ability to monitor and modify speech. Therefore, students learning English for international communication must learn to speak it intelligibly and comprehensibly, not necessarily like natives but well enough to be understood (Gilakjani, 2012).

Drilling is listening to a model and repeating what the model said. It is a repetition drill, which many teachers still use when introducing new language items to students. The teacher speaks the word, and the students repeat it (Derwing & Munro, 2015). It is not a new or fashionable classroom technique, but it can be of great value to the learners if used appropriately in the classroom. Avoid over-drilling and keep the stages alive. Respond to the needs of the learners and drill if it will help them pronounce or memorize words or language chunks. To help students remember the language, vary how drills are conducted (Sewell, 2016).

There are drilling activities a teacher can use to improve language acquisition. The most common is Repetition drills. Learners listen to the model and then repeat what the model said. Words should be clear, natural-sounding, and consistent. The drill varied concerning who repeats-whole class, half the class, boys only, girls only, and individuals (Sewell, 2016; Agudo, 2014; Derwing & Munro, 2015).

*Theoretical Framework.* This study anchored on the Behaviorist theory that learning to speak a foreign language was a matter of correct habit formation. It was thought that the

correct repetition of phrases often would lead to the mastery of the language. One of the tenets of Behaviorist theory is the habit formation of language teaching and learning. Language learning is not problem-solving but the information and performance of habits (Johnson, 2017). Language learning is a mechanical system that leads learners to develop patterns, the fundamental scheme of a conditioned reflex. *L*anguage learning and its development is a matter of conditioning through imitation, practice, reinforcement, and habituation, which constitute the paces of language acquisition (Moeller & Catalano, 2015).

This study also adhered to the Skill Acquisition Theory. It is a learning theory that ranges from cognitive to psychomotor skills. This theory claimed that learners commence learning something through mainly explicit processes and proceed to implicit methods through subsequent sufficient practice and exposure. Second-language acquisition is learned in the same way as any other skill, such as driving a car or playing the piano. They see the practice as the vital ingredient of language acquisition (VanPatten & Williams, 2008).

**Methodology**

*Research Design.* This study used an action research design. Action research is a participatory, democratic process that seeks together action and reflection, theory and practice, participation with others to pursue practical solutions to pressing concern issues to people (Efron & Ravid, 2013). The researcher determined the learners' pronunciation skills; and identified oral drills as an appropriate intervention to improve the skill. After ten cycles of oral drills, a post-test was done to check if the said skills improved.

*Research Setting.* The research was conducted in Aloran Central School, located in Aloran, Misamis Occidental, Northern Mindanao, Philippines. The school has 36 teachers from Kindergarten to Grade 6. Each grade level has an average of 3 sections, and each section has an average of 35 learners. Since it is the central school of Aloran, it caters to the villages' learners in the town center area.

*Participants.* The research involved the Grade 6-Earth pupils of Aloran Central School for SY 2018-2019.  There were 34 participants, 20 males, and 14 females. The researcher selected the participants through convenience sampling because they were all in the researcher's advisory class.

*Research Instruments.* These were the research instruments used in the study:

> A. Pronunciation Skills Checklist (Pretest/ Posttest)
> During the conduct of The Philippine Informal Reading Inventory (PHIL IRI) at the onset of the school year, the researcher identified the most common sounds that the learners have difficulty pronouncing. The top 10 consonant sounds were /f/, /hw/, /v/, /sh/, /th/, /th/, /nt/, /z/, /k/, and /zh/. The top 10 vowel sounds were /a/, /ô/, /ou/, /ā/, /i/, /e/, /oi/, /ô/, /ō/, and /o͞o/. The pretest used ten words for the consonant sounds category and another ten words for the vowel sounds category. The posttest utilized another set of words.

In determining the learner's level of pronunciation skills during the pretest and post-test, the researcher used the following scale:

| | |
|---|---|
| 9-10 | Excellent |
| 7-8 | Very Satisfactory |
| 5-6 | Satisfactory |
| 3-4 | Less Satisfactory |
| 0-2 | Unsatisfactory |

B. Oral Drills Word List
The researcher gathered words per sound and was used in phrases and sentences. These sounds were the focus of the oral drills done in 10 weeks. The focus of the drill was one consonant and one vowel sounds per week.

C. Interview Questions
This study utilized interview guide questions which contained six items. The first two questions focused on the learner's assessment of one's pronunciation skills. The second two questions emphasized the learner's personal views of the oral drills. Furthermore, the last two items stressed the learner's assessment of pronunciation skills after undergoing oral drills.

*Data Collection.* At the school year's onset, the researcher identified the least pronounced sounds based on the PHIL IRI materials; ten consonant sounds and ten vowel sounds. The pupils were given a list of words in determining their pretest performance. Two English teachers served as co-raters who helped establish the pronunciation skills. After the pretest, the oral drills followed. The oral drills focused on the 20 sounds that the learners have problems pronouncing. The researcher read the words first, then the pupils repeated them. The drills transitioned from words to phrases and sentences, which lasted for ten weeks. It focused on one consonant and one vowel sounds per week. It utilized the reading period of the class, which lasted for 30 minutes. Day 1 focused on a consonant sound, and the whole class repeated the words. Then individual learners repeated the same words on the second day. On the third day, the drill transitioned to a vowel sound as a whole class and individually on the fourth day. Moreover, on the fifth day, individual learners pronounced ten words covered in the week. The drilling lasted for ten weeks covering all 20 sounds. After completing the last cycle, the researcher conducted a post-test using a different set of words to determine if the pupils retained and could produce specific sounds correctly. The same two English teachers who served as co-raters during the pretest helped assess the pupils' post-test performance.

A semi-structured interview was conducted with the participants. Each interview was done in the vernacular to ensure that the participants understand the questions and answer them. The conversation ranged in length from approximately 10 to 20 minutes and was recorded using a tape recorder. After the interview, the answers were transcribed verbatim, analyzed, defined, and coded. The researcher followed the protocol in an interview during the conduct of the in-depth interview.

*Data Analysis.* After the pretest and post-test, the number of correctly pronounced words were counted, and scores were categorized as Excellent: 9-10, Very Satisfactory: 7-8, Satisfactory: 5-6, Less Satisfactory: 3-4, and Unsatisfactory: 0-2. The frequency determined the number of participants in a particular category.  The percentage determined the proportion of the participants in the distribution based on their performance. The mean was determined after combining all participants' scores for the pretest and post-test, and the standard deviation was the value showing deviation from the mean. T-test was used to test the significant difference in the learner's pronunciation skills in the pretest and post-test. The transcripts of the interview were analyzed through thematic analysis.

**Results and Discussion**

*Level of Learner's Pronunciation Skills in the Pretest.* Table 1 shows a summary of the level of learner's pronunciation skills in the pretest. In pronouncing consonant sounds, 53% of the learners got satisfactory, while 29% got very satisfactory. In pronouncing vowel sounds, 44% of the learners got satisfactory while 50% got very satisfactory. The majority of the learners got scores of 5 and 6 in pronouncing consonant sounds. While most of them got scores of 7 and 8 in pronouncing vowel sounds, this indicated that while they pronounced most of the sounds, they still struggled to pronounce some sounds.

*Table 1: Summary of Learners' Pretest Performance*

| Sounds | Performance | Frequency | Percentage |
|---|---|---|---|
| Consonants | Excellent | - | - |
|  | Very Satisfactory | 10 | 29.00 |
|  | Satisfactory | 18 | 53.00 |
|  | Less Satisfactory | 5 | 15.00 |
|  | Unsatisfactory | 1 | 3.00 |
|  | Total | 34 | 100.00 |
| Vowels | Excellent | - | - |
|  | Very Satisfactory | 17 | 50.00 |
|  | Satisfactory | 15 | 44.00 |
|  | Less Satisfactory | 2 | 6.00 |
|  | Unsatisfactory | - | - |
|  | Total | 34 | 100.00 |

More learners can pronounce vowel sounds correctly than consonant sounds. The majority of students had difficulty pronouncing consonant sounds. They had problems pronouncing the "th" sound. It is one of the most challenging consonant sounds to pronounce. It can be pronounced in three different ways: as a "d" as in this, that, these, those, they or them; as the voiceless in three, thing, thought; or as a /t/ as in Thai or Thames. The voiceless t is especially tricky for some students, often saying tree instead of three (Pesce, 2017).

Some learners also struggled in pronouncing vowel sounds. Vowels have five letters, which are a, e, i, o, u. They map to 13 different sounds; for instance, several sounds of the letter o in words: boat, boot, out, and hot. Pronouncing vowels is one of the problems that Asian adult learners encountered in pronunciation subjects. They have difficulty with the /a/ sound because hearing and pronouncing the sound is difficult. They cannot differentiate between the long sound /a/ and the sound /e/. They also have problems distinguishing between the long /e/ and the short /i/ (Long, 2011).

The results of the thematic analysis of the interviews have validated learners' inability to pronounce some sounds. Based on the findings of the interview, the participants exhibited apprehension. They were apprehensive because their classmates would laugh at them if they could not pronounce the words correctly.

The researcher's notes also validated this. Based on the notes, the learners had difficulty pronouncing the words because they did not know how these words were pronounced. They were also afraid that others would laugh at them if unable to pronounce the words. This anxious feeling can be considered performance anxiety that belongs to social anxiety (Whiting et al., 2015). It involved an individual's fear and worry of being perceived and evaluated negatively by others. External cues and situational demands that include concerns about other's evaluations of one's behavior triggered this.  Performance anxiety had contexts in which individuals were exposed to possible external negative evaluation by others. Second language learning is an example of performance anxiety. In the second language acquisition context, teacher's and peers' assessment of second language learners made them feel apprehensive. Communication apprehension is observed in oral communication contexts and is defined as the level of anxiety of a person brought about by either real or anticipated communication with others (Spinner & Gass, 2014).
The learners were apprehensive because they did not know the phonemes of the words. Phonemic awareness is the ability to hear and operate the sounds in spoken words and understand that speech sounds comprise spoken words and syllables. It is essential to learning to read in an alphabetic writing system because letters represent sounds or phonemes.

Another struggle was that some of these words were new to them. They might have encountered these before, but they did not know how they are pronounced correctly. Since they did not know how to pronounce them, they did it based on the spelling. Some learners read a word as it was spelled. It was often hard to know how to pronounce a word by its spelling. Sounds are the basis of English pronunciation, not spelling. Error in these areas can

make it difficult to be understood by native English speakers. The spelling and their pronunciation lack correspondence.

There are several ways to pronounce a particular spelling pattern, but it certainly helps to know the variations. Many words in the English language have silent letters, but these words are unknowingly pronounced, making these words mispronounced. There are numerous letters and combinations of letters which produce various sounds at different places (Nawaz, 2011).

In the class, both consonant and vowel sounds should be emphasized because both sounds are essential. Teachers should provide an environment that does not add anxiety to the learners. If learners cannot pronounce some words, they should be corrected in a positive tone to have a positive attitude towards language learning. Other learners should be discouraged from making fun of someone's mistakes because it is normal to commit errors.

*Level of Learner's Pronunciation Skills in the Posttest.* Table 2 shows a summary of the level of learner's pronunciation skills in the post-test. In pronouncing consonant sounds, 74% of the learners got very satisfactory, while 24% got excellent. In pronouncing vowel sounds, 71% of the learners got very satisfactory, while 29% got excellent. The majority of the learners got scores of 7 and 8 in pronouncing both sounds. There was an increase in the number of learners who got an excellent score. Their level of pronunciation skills increased from satisfactory to very satisfactory. Their performance in the post-test improved from the pretest after ten cycles of oral drills.

*Table 2: Summary of Learners' Posttest Performance*

| Sounds | Performance | Frequency | Percentage |
|---|---|---|---|
| Consonants | Excellent | 8 | 24.00 |
| | Very Satisfactory | 25 | 74.00 |
| | Satisfactory | 1 | 3.00 |
| | Less Satisfactory | - | - |
| | Unsatisfactory | - | - |
| | Total | 34 | 100.00 |
| | | | |
| Vowels | Excellent | 10 | 29 |
| | Very Satisfactory | 24 | 71 |
| | Satisfactory | - | - |
| | Less Satisfactory | - | - |

| | | |
|---|---|---|
| Unsatisfactory | - | - |
| Total | 34 | 100.00 |

The sounds covered in the oral drills were given two days of practice to provide enough time for the learners to master the said sound. It gave the learners enough time to listen and practice as a group and individually. While improving pronunciation might seem unattainable, helping students improve their pronunciation one sound at a time is much more doable. Instead of taking up most class time practicing pronunciation, practice a different phoneme every day or every week (Pesce, 2017).

The thematic analysis of the interview validated the learner's improvement from pretest to post-test. Based on the findings, the participants found the oral drills useful in improving their pronunciation skills. Second-language acquisition is learned in the same way as any other skill. The practice is the vital ingredient of language acquisition (VanPatten & Williams, 2008). They were inspired to practice more because they liked doing it. It helped them pronounce words correctly. It gave them time to exercise their mouth and tongue. They were able to pronounce words more precisely because they knew how. It was their only venue to practice pronouncing words because this skill was not given enough time in the English subject. The English curriculum is content based; there is no emphasis on pronunciation. Their English teacher taught them pronunciation, but it was not the focus of the lesson. The content of the English subject in the 6th grade is no longer focused on pronunciation. The researcher is not an English teacher; he is a science teacher. However, he has a background in English because he used to be a call center trainer. The oral drills were done during the remedial period; the participants were given enough time to learn the pronunciation and practice how to do it.

The researcher's notes also validated this. Based on the notes, the learners were participative during the drilling. They liked doing it. There was a struggle in the first few cycles, but they have eventually improved in succeeding ones. They looked forward to doing the oral drills. They even practiced pronouncing the words even after the allotted time. Drilling helped learners memorize language through the teacher's control, where students' mistakes can be corrected and encouraged to fix difficulties simultaneously. For the learners, drills provided a safe environment to experiment with producing the language. It may help build confidence. It helped students notice the correct form or pronunciation of a word or phrase. Consciousness-raising of language is an important stage in developing language competence. Teachers should give learners immediate feedback on their accuracy concerning. Many learners want to be corrected. It helped memorize and automate common language patterns and language chunks (Sewell, 2016; Agudo, 2014; Derwing & Munro, 2015).

In the class, learners eventually learn if given enough time to practice. They will be able to pronounce words correctly if given time to do it. Teachers at the lower grade levels should provide ample time for teaching pronunciation. However, there are many skills one should learn; pronunciation is just one of them.

*Test of Significant Difference between the Level of the Learners' Pronunciation Skills in the Pretest and Posttest.* In determining the significant difference in the level of learners' pronunciation skills between the pretest and post-test, the researcher used t-tests, mean and standard deviation. Table 3 presents the data.

Table 3 presents the summary of the pretest and post-test of the consonant and vowel sounds. The learners' mean score in the pretest for consonant sounds is 5.6, while the post-test is 7.9. The mean of the pretest for the vowel sounds is 6.3, while the post-test is 8.2.

*Table 3: Test of Significant Difference in the Level of the Learner's Pronunciation Skills in the Pretest and Posttest*

| Variables | Mean | SD | Computed T-test | Critical Value ($\alpha = 0.05$) | Remarks |
|---|---|---|---|---|---|
| Consonants | | | | | |
| Pretest | 5.6 | 1.328 | -8.16007 | .00001 | There is a significant difference. |
| Posttest | 7.9 | 0.995 | | | |
| Vowels | | | | | |
| Pretest | 6.3 | 1.069 | -7.4368 | .00001 | There is a significant difference. |
| Posttest | 8.2 | 1.016 | | | |

For the learners' performance in the pronunciation of consonant sounds, the significant difference test yielded the t-value of -8.16007, which is less than the critical value of .00001 at a .05 percent level of probability. There is an indication that the learners' performance on the post-test is significantly different from the pretest. For the learners' performance in the pronunciation of vowel sounds, the significant difference test yielded the t-value of -7.4368, which is less than the critical value of .00001 at a .05 percent level of probability. There is an indication that the learners' performance in the post-test is significantly different from the pretest.

The oral drills done in 10 cycles brought these improvements in the learners' pronunciation of consonant and vowel sounds. The constant drilling helped them master the sounds.

For the learners, drills provided an emphasis on accuracy. Increased accuracy, fluency, and complexity improve the learner's language. There is a need to emphasize accuracy at certain stages of the lesson or during certain task types. It provided learners with intensive practice in hearing and saying particular words or phrases. It helped learners get their tongues around difficult sounds or imitate intonation that may be slightly different from their first language (Sewell, 2016; Agudo, 2014; Derwing & Munro, 2015).

The thematic analysis of the interview validated the learner's improvement from pretest to post-test. Based on the findings, the participants thought that having good pronunciation skills would make them feel good. People look up to them, and they would not laugh at them because they can pronounce English correctly.

The researcher's notes also validated this. Based on the notes, the learners believed that pronouncing words will make them feel better. It gave them a sense of fulfillment because no one would laugh at them. It was a typical experience that every time they mispronounced a word, their classmates laughed at them. When they moved to higher grade levels or even college, they could pronounce words correctly. Moreover, when they worked, they would perform their job well because they communicate in English fluently. According to Linge (2014), we all know that we should not judge a book by its covers, but we still do, unconsciously, most of the time. People tend to underestimate people who have lousy pronunciation and overestimate people who have the proper pronunciation. For instance, think about immigrants in a foreign country who speak a broken version of the native language. Even though we do not want to, it is easy to think that foreigners with good pronunciation are "better" than those with poor pronunciation. Pronunciation, unlike the other skills, strikes the listener directly in the face. How good one's pronunciation is, in general, can be judged very quickly, and an opinion is formed automatically by anyone who hears the person. Students achieve better results by being aware of the importance of pronunciation and motivation for practicing it. The student's disposition to pronunciation was the key predictor in the learning of native or near-native pronunciation. The more concerned they were, the better was their quality.

We live in a world where English is an international lingua franca, where many job positions presently require excellent English knowledge. If students do not acquire proper pronunciation, they will face difficulties finding employment (Gilakjani, 2012).
In the class, learners think that they need to be skillful in English to have better future job opportunities. If they do not acquire proper pronunciation, they will face difficulties in finding employment. With the Philippines' BPO industry boom, some learners look forward to working in this English-speaking job.

**Conclusion**

The learners' pronunciation skills significantly improved from the pretest to the post-test after oral drills for ten cycles. The oral drills introduced to the learners improved their pronunciation skills from Satisfactory to Very Satisfactory. The constant drilling worked. It

helped them develop the pronunciation of the least pronounced consonant and vowel sounds.

The researcher recommends that this intervention be adopted in the lower grade levels, focusing on all sounds so that learners are already familiar with the phonemes when they move to higher grade levels. The researcher also recommends that another action research is done to improve comprehension skills.

**About the Author**

**Antonio Jr. T. Enerio, PhD** is a graduate of Doctor of Philosophy in Education major in Educational Management at La Salle University, Ozamiz City.  He is the OIC Dean of the School of Teacher Education in Northwestern Mindanao State College of Science and Technology, Tangub City. He is currently teaching Assessment of Learning to teacher-education students. He used to be a teacher in the Department of Education before moving to a state college. His research interests include action research, student engagement, and assessment. Email: antoniojr.enerio@nmsc.edu.ph

## References

Agudo, J. (2014). English as a foreign language teacher education: Current perspectives and challenges. Rodopi.

Beare, K. (2017). Understanding English Pronunciation Concepts. https://www.thoughtco.com/understanding-english-pronunciation-concepts-121

Derwing, T., & Munro, M. (2015). Pronunciation Fundamentals: Evidence-based perspectives for L2 teaching and research. John Benjamins Publishing Company.

Efron, S., & Ravid, R. (2013). Action Research in Education: A Practical Guide. Guilford Press. Canadian Journal of Program Evaluation. https://doi.org/10.3138/cjpe.30.1.103

Gilakjani, A. (2012). A Study of Factors Affecting EFL Learners' English Pronunciation Learning and the Strategies for Instruction. International Journal of Humanities and Social Science 2 (3). http://www.ijhssnet.com/view.php?u=http://www.ijhssnet.com/journals/Vol_2_No_3_February_2012/17.pdf

Johnson, K. (2017). An Introduction to Foreign Language Learning and Teaching. Routledge.

Linge, O. (2014). Two reasons why pronunciation matters more than you think. https://www.hackingchinese.com/two-reasons-why-pronunciation-matters-more-than-you-think/.

Long, M. (2011). Problems in Second Language Acquisition. Routledge.

Moeller, A., & Catalano, T. (2015). Foreign Language Teaching and Learning. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1195&context=teachlearnfacpub.

Nawaz, A. (2011). Problems faced in English Pronunciation. Retrieved from http://sunnyenglishliterature.blogspot.com/2011/05/problems-faced-in-englishpronunciation.html.

Nordquist, R. (2016). pronunciation (English): Glossary of Grammatical and Rhetorical Terms. https://www.thoughtco.com/pronunciation-english-1691686.

Nordquist, R. (2017). What is Communication? The Art of Communication - Verbal and Nonverbal. https://www.thoughtco.com/what-is-communication-1689877.

Pesce, C. (2017). 7 Worst Pronunciation Mistakes ESL Students Make Around the World. https://busyteacher.org/14846-7-worst-pronunciation-mistakes-esl-students-make.html

Sewell, A. (2016). English pronunciation models in a globalized world: Accent, acceptability and Hong Kong English. Journal of Second Language Pronunciation. https://doi.org/10.1075/jslp.19034.thi

Spinner, P., & Gass, S. (2014). Second language acquisition and linguistics: A bidirectional perspective. Linguistics Vanguard 1(1). https://doi.org/10.1515/lingvan-2014-1010

Smith, M.S., & Candlin, C. (2014). Second Language Learning: Theoretical Foundations. Routledge.

VanPatten, B., & Williams, J. (Eds.) (2008). Theories in Second Language Acquisition: An Introduction. *Studies in Second Language Acquisition 30(02).* https://doi.org/10.1017/S0272263108080418

Whiting, P. F., Wolff, R. F., Deshpande, S., Di Nisio, M., Duffy, S., Hernandez, A. V., Keurentjes, J. C., Lang, S., Misso, K., Ryder, S., Schmidlkofer, S., Westwood, M., & Kleijnen, J. (2015). Cannabinoids for Medical Use: A Systematic Review and Meta-analysis. JAMA, 313(24), 2456–2473. https://doi.org/10.1001/jama.2015.6358

# INCORPORATING AUTONOMY IN AN ANALYTICS MODULE: VISUALIZING SELF-DIRECTED LEARNING

**Hui Teng Chia**
**Singapore Polytechnic**

**Abstract** To instill self-directed learning (SDL) in post-secondary learners, there has to be a transfer of the responsibility of learning to the learners themselves.  However, a transfer of autonomy in learning from the teacher to the learners may not always be feasible especially in traditional classroom setting with tight teaching schedules and pre-defined syllabus to cover.  Furthermore, post-secondary learners may be perceived to yet have the maturity to decide on what they want to learn, how they learn it and to evaluate their own learning.  This study examines the impact of an increase in autonomy of learning on learners' conception of SDL.  A group of 40 participants enrolled in a visual analytics module are asked to decide on a learning aspect, set learning goals and evaluation criteria to evaluate their own learning.  This study offers two visualizations to illustrate how learners make sense of SDL in a minimalist autonomy learning environment and conclude with some thoughts on the role of autonomy in facilitating self-directedness in learners.

## Introduction

The economic uncertainty due to the pandemic, as well as the rapid acceleration of automation, has prompted a growing number of people making career switch to new fields (Russo, 2020).  Such transition requires workers to have the capability in identifying their own knowledge gaps and to close these gaps to meet the demands of new careers. Hence, it is critical for pre-employment training of post-secondary learners for such workforce to be equipped with learning agility to upskill and reskill.  This includes the necessary thinking tools on how to diagnose their own learning needs, set learning goals and conduct self-evaluation on their own learning.  In short, post-secondary learners need to know how to be self-directed learners. One of the essential components to instil self-directedness is autonomy.  It involves the transfer of learning responsibility from the teacher to the learner. While the literature does not lack in research on self-directed learning, there are limited studies in showing the process between the extent of autonomy in learning and the development of self-directedness.  This study sets out to find the answers for two aspects of autonomy and self-directed learning, which are, learner readiness in embracing the transfer of responsibility of learning to themselves, and the indicators of self-directedness in the process of learning.

## Literature Review

*Self-Directed Learning and Autonomy.*  A self-directed learner is an attribute that every educator aspires for persons schooled in the Singapore formal learning institutions (MOE Singapore, 2015). This is because self-directed learners are self-actuated learners taking responsibility for their own learning that moves them from their current states to where they want to be (Piskurich, 1993).  Learners who are self-directed are seen to be "taking initiative, diagnose their learning needs, formulate learning goals, identify human and material resources for learning, choose and implement appropriate strategies and finally evaluating their learning outcomes" (Knowles, 1975, p.18).

The notion of SDL has its roots in adult education which began to receive much attention in the 1960s (Knowles, 1980).  Adults are observed to prefer more self-directedness in learning in which a majority of all their learning projects are planned out by the learners themselves (Tough, 1971, as cited in Knowles, 1973).  Such preference for independence is associated with adults' accumulation of life experiences and their needs to solve problems in their social roles such as in their jobs.  Children and youths are seen to be lacking such life experiences and are missing the responsibilities of social roles.  Therefore, children and youths are deemed to require much dependency on external support like a teacher, to plan, manage, monitor and evaluate their learning.

However, Knowles (1980) argued that youths especially do start early in life to accumulate experiences that have values for learning.  Such observation can be seen in youths taking on social roles like part-time jobs, internships as well as taking on responsibility in managing their own lives.  Children too could show signs of self-directedness in one area, such as learning a new game, but may show dependency in other areas such as when it comes to deciding how to learn a skill.  Hence, rather than seeing self-directedness as a dichotomous entity, it should be viewed in a situation-dependent continuum entity regardless of age maturation (Knowles, 1980).

There seems to be unanimous agreement amongst researchers on the "freedom and the agency of the learner as steward of creation" (van der Walt, 2019, p.1).  That is, whether SDL is associated with the characteristics of a learner (e.g., Douglas & Morris, 2014; MOE Singapore, 2015), the process or an approach to learning (e.g., Bartholomew, Reeve, Vion, Goodridge, Lee & Nadelson, 2017; Knowles, 1980; Peine, Kabino & Sprecklesen, 2016) or as training designs (e.g., Piskurich, 1993, Gibbons, 2002), the emphasis on independence of the self is strong.  As such, the notion of autonomy is central in SDL.

Ryoo (2011) describes the notion of autonomy from four different perspectives - origin, directionality, boundary, and treatment.  In terms of origin, autonomy can be initiated by the self (*autogenic*) or prompted by an external stimulus (*heterogenic*).  For example, a person who wants to be an entrepreneur due to his or her own desire displays autogenic autonomy, and a person who wants to be an entrepreneur due to persuasion of the society displays heterogenic autonomy.  In terms of directionality, autonomy resulted in freedom from external hindrance is known as *negative* autonomy.  In comparison, autonomy resulted in freedom to exercise internal capacity to fulfil one's freedom is known as *positive*

autonomy.  For example, a person with no admission restrictions to choose from different entrepreneurial courses has negative autonomy, whereas a person who knows the rigor of each entrepreneurial courses has positive autonomy in choosing.

Ryoo describes the boundary of an autonomy as *strong* when it strives for an outcome beyond socially, traditionally, and culturally defined parameters.  On the other hand, *weak* autonomy seeks to preserve independence within such parameters.  For example, a learner who proposes a new objective, deadline, and rubrics to replace an existing assessment shows strong autonomy, whereas a student who decides not to complete an existing assignment shows weak autonomy.  Finally, in terms of treatment, autonomy can be *descriptive* or *ascriptive*.  Descriptive autonomy involves suppressing a person's own freedom to decide on the next course of action due to perceived incapacity of this person in making a good decision.  Ascriptive autonomy involves the respect for a person's free will to decide even though this person is perceived to be incapable of making a good decision. An example of descriptive autonomy includes imposing what a learner needs to learn in an entrepreneurial course, and in contrast, a course that allows learner to decide what modules to take is exercising ascriptive autonomy.

Taken altogether, Ryoo suggests that there exist two views each at one end of a spectrum of autonomy - maximalist and minimalist.  The maximalist view subscribes to autogenic, negative, strong and ascriptive autonomy, whereas the minimalist view adopts the heterogenic, positive, weak and descriptive autonomy.  It is tempting to perceive that learners will be motivated to learn when they are given the autonomy to decide what they want to learn, how they want to learn it and how they would like to be assessed in their learning (Balser, 2018), but Ryoo argues that instead of an absolute decision of all or no autonomy, the practical approach in education is to adopt the stance to offer varying degree of autonomy transfer to learners depending on the levels of education.

Although SDL is a natural psychological progression as people mature (Knowles, 1980), there may be resistance amongst learners (even adult learners) when the responsibility of teaching is transferred to the learners to encourage autonomy.  SDL is no easy feat for many learners especially for those who are conditioned by their previous experiences to have dependency on external influence in learning (Knowles, 1980), those with low self-esteem or those who had experienced a series of failures in their learning journey (Gibbons, 2002).  Such responsibility to take charge of one's own learning may also be overwhelming even for capable learners as the responsibility to keep track of learning starts to accumulate (Gibbons, 2002).

**Methodology**

*Research Questions.*  This study examines how the increase in autonomy is associated with learners' engagement in SDL.  Specifically, the research questions for this study are as follows:
  1. To what extend did the participants in this study embrace (or resist) an increase in autonomy of learning?

2. What are the differences in how learners make sense of SDL in a module designed using Ryoo's (2011) minimalist view on autonomy?

*Participants.*  Forty second-year full-time diploma learners between the ages of 18 to 23 years participated in this study.  The participants, comprising 23 males and 17 females, are enrolled in a visual analytics diploma module.  The lessons are conducted in 15 weeks, four hours per week from which an hour is scheduled for the participants to learn on their own from pre-recorded video lectures.  In the three hours face-to-face lessons, the participants solved scenario-based visual analytics problems by generating visualizations using the *Tableau* software.  The content coverage in this module includes introduction to the analytics thinking process, the different types and purposes of visualizations as well as best practices in building dashboard and storyboard.  Participants need to complete five assessment components designed to include some degree of autonomy in learning.  The weightages of these assessment components range from 15% to 35% of the overall grade in the module.  Due to space limitations, only one of the five assessment components will be discussed in this paper. Henceforth, this assessment component would be referred to as assignment.  At the end of the semester, participants are required to complete a reflection survey on their perceptions of SDL.  Twenty-six participants completed all the required components of the assignment. Fourteen participants did not complete the reflection survey in which eight of them also did not complete the assignment.  All these data are consolidated to form the data for the main study.

*Research Instrument.*  The aim of the assignment is for the participants to explore areas of their interest and to extend their learning beyond the content taught in this visual analytics module.  Specifically, participants are given the autonomy to decide what they want to learn, how they wish for the learning to occur and at the end, they would do a reflection on their learning.  Generally, this assignment allows for controlled exploration within the pre-defined parameters adopted from Ryoo's (2011) minimalist view on autonomy - heterogenic, positive, weak, and descriptive.  Since the onset of exploration for this assignment is external which is to fulfil the requirement of this module, the autonomy is heterogenic in nature.  The assignment allows for participants to exercise positive autonomy in determining how they want to fulfil the requirement of this assignment.  Due to administrative constraints, the assignment offers weak autonomy as it requires participants to work within the pre-determined structure and does not encourage participants to change the weightage and deadline set. In addition, the assignment includes descriptive autonomy as it requires participants to trust that by completing the assignment as designed, they would attain a greater educational goal in return for suppressing their own personal desires for the time being.

This minimalist autonomy designed assignment is divided into three phases.  Participants' responses in each phase are gathered as evidence of their SDL.

*Phase 1: Goal setting.* At the start of the semester, the participants are briefed on the aim of the assignment.  After the participants clarify the scope and deliverables of the assignment, they proceed to individual goal setting.  In this first phase, the participants describe their

areas of interest and set own learning goals.  Then the participants set learning milestones and assessment rubric to evaluate their own learning.

*Phase 2: Monitoring own learning.* Following the completion of Phase 1, participants are encouraged to keep a weekly e-journal in a learning management system to monitor their own learning and to document all evidence of milestone achievements.  In this Phase 2, the participants are encouraged to communicate with their tutor regularly to get feedback on their learning.

*Phase 3: Evaluating own learning.* The final part of this assignment is the evaluation process where participants meet with their tutor to reflect on the goals set, their self-monitoring of the learning process and to perform a self-assessment.  Finally, participants complete a reflection survey as a signal for the completion of this assignment.

On average, participants are given about 10-12 weeks to complete Phases 1 to 3.  Figure 1 maps the overall roll-out of this assignment against the academic term of this module which begins in mid-April and ends by mid-August.

| Academic calendar | April | | May | | June | | July | August | |
|---|---|---|---|---|---|---|---|---|---|
| | | Term | | Exam | Term break | | Term | August Exam | |
| This study | | PHASE 1 Start of Assignment (Briefing and goal setting) | PHASE 2 Assignment in Progress (Monitoring own learning) | | | | | PHASE 3 End of Assignment (Evaluation of learning) | |
| | | | Reminder 1 | | Reminder 2 | | Reminder 3 | | |

Figure 1. The implementation timeline

*Data collection and analysis.* There is a two-fold interest in this study.  First, I would like to know if participants embrace or resist a transfer of the responsibility of learning from the teacher to the learners themselves.  This would inform me of their readiness for more of such activities in subsequent design of curricular content.  Second, I am interested in the process of how learners engage in a more autonomous way of learning.  Participants are deemed to resist the increase in autonomy when they do not participate in this assignment or, they express anxiety in learning on their own.  Participants who embrace the autonomy would set goals.  What they set as goals, as well as their responses to the survey questions would uncover what values in this assignment that are consistent with their own values in learning.  It would also shed light on whether the participants adopt positive stance towards such an increased in autonomy, or they face a lot of anxiety.

All the data in Phases 1 and 2 of the assignment are collected and stored in an online LMS.  These data are part of participants' course work requirements.  Survey data collected via Google Form in Phase 3 are similar to feedback that are routinely asked of learners for module delivery improvement.  Hence, participants are not disadvantaged in any way by the data collection.  The data are compiled into a single file and all sensitive information are de-identified for further analysis.  Table 1 summarizes the data analysis methods on data collected to answer the research questions.

*Table 1: A summary of research questions, data collection and the data analysis methods.*

| Research questions | Data collection | Data analysis | |
|---|---|---|---|
| **To what extend did the participants in this study resist or embrace an increase in autonomy of learning?** | Number of participants who submitted or did not submit the assignment. | Descriptive statistics | The responses gathered are analyzed per variable and then analyzed per participant to give rise to the different profile of self-directed learners in a learning environment designed with minimalist view on autonomy. |
| | Responses from goal setting *What do I want to learn?* | Qualitative coding | |
| | Responses from the survey questions *Which part of this assignment makes sense to you?* | | |
| **What are the differences in how learners make sense of SDL in a module designed using the minimalist view on autonomy?** | Lapses (in days) between goal setting and the first e-journal submission | Descriptive statistics | |
| | Number of e-journals submitted | | |
| | Median days between e-journal submissions. | | |
| | E-journal submissions on whether the reflection of learning was directed towards achieving the goals set. | Qualitative coding | |
| | Responses from the survey questions *When did you feel a sense of success when you did this assignment?* | | |
| | Responses from survey questions *To me, self-directed learning is _____.* | | |

Quantitative responses such as number of submissions and the date of submissions collected from LMS are analyzed and reported using descriptive statistics.  Subsequently, open-ended responses from the e-journal submissions and final survey are analyzed qualitatively.  The e-journal submissions are analyzed for evidence on whether the participants constantly reflected on how their efforts are directed at achieving the goals set. The e-journal submissions with strong evidence on effort directed at achieving goals are coded as *Monitor-Focused*.  On the other hand, responses are coded as *Monitor-Distracted* when there is evidence of effort but not directed at goals, or with no evidence of effort identified. Responses from the survey question - *When did you feel a sense of success when you did this assignment?* are coded as *Success-Completion* when the sense of success is derived from completing the task, as *Success-Beliefs* when the sense of success came from a held belief that is changed, and *No-Evidence* when there is no evidence on the sense of success in which reasons will be discussed instead.  Finally, each participant's conception of SDL is derived from his or her response to the survey question *To me, self-directed learning is ____.*  Responses are coded as *Holistic-SDL* when there is an emphasis of self, with or without external help, in goal-setting, monitoring and evaluating own learning.  When some

of these elements are mentioned, the responses are coded as *Partial-SDL*.  If none of the SDL components are mentioned, then the responses are coded as *No-Evidence*.

The following section will first discuss the evidence gathered to answer the first research question, and subsequently, the discussion will focus on answering the second research question.

**Results and Discussion**

*To what extend did the participants in this study resist or embrace an increase in autonomy of learning?* Ten out of 40, or 25% of the participants showed some levels of resistance towards an increased in autonomy of learning.  Two of these participants expressed their apprehension to take on additional learning on their own.  Instead, they expressed their desire to master the basic skills taught in class.  This shows that they are not yet ready to explore learning on their own beyond what is taught in the classroom and needed more guidance from an authority.  An example of such is given by participant S36: "I would like to learn the basic skills on [software] as I am still very unsure on where to put the datas from Measure to the CARD, I need extra guidance for the basic skills so that i can go further ahead to use the skills without having any difficulties."

Eight other participants hinted subtle resistance towards an increase in autonomy of learning by not completing the assignment.  These participants chose not to complete this assignment despite repeated reminders to complete it, and that the assignment carried significant weightage to the overall grade received for this module.   The lack of participation in this assignment hinted at their unwillingness to devote time in setting own learning goals, monitor their own learning and to evaluate their learning.  It is possible that participants may be apprehensive towards such an unfamiliar assessment structure that transfer the responsibility of learning from the teacher to the learners themselves.  Such apprehension could be due to previous schooling experiences have conditioned them to expect structure and dependency on a figure of authority (Knowles, 1980) or lack of self-confidence that they could direct and validate their own learning (Gibbons, 2002).  Besides that, it could also hint at their beliefs that such assignments with increased autonomy generate little values and are indulgence for those who have time to explore learning on their own.  As such, when there are other tasks that compete for their limited time and attention, it is possible that those that they place higher values than this assignment would capture their attention and interest first.  Participant S38 offers a glimpse of such a reason: "After the first stage, which is to collect my data for my "Laughter" project, i became a little busy and occupied by my other modules and completely lost track of time (…) It happened when i had projects and assignments from other modules. Therefor i have decided to focus on the "more important" assignments as i had dateline to catch."

The remaining 30 out of 40, or 75% of the participants seemed to embrace an increase in autonomy of learning.  The goals that these participants set were technical in nature and related to the module curriculum content.  These goals mentioned the desire to explore different types of visualizations, animation in visuals and the symbols and signs in the

visualizations, as indicated by S29 and S18 respectively.  Others described how they want to apply what they have learnt in an area that interests them, such as an example given by S3.

- S29: " (…) after watching Hans Rosling Ted Talk video, i have learned that how visuals and colors and animation play a big part in data visualization. I have watched the video about him explaining the data about child fertility rate and life expectancy, i am keened to learn how to use visuals and colors as well as animation to help myself and others understand the dat[a] that I have been give[n] a bit more."
- S18: "Able to learn at least 3 new charts which are not taught within our module. Would like to learn them as i get to explore more choices of charts when doing future analytics."
- S03: "I would like to apply the knowledge and skills that I've learnt in class and use them to present data related to one of the sports in which I am rather interested in, Formula One (…) data in a visual way would be rather interesting and may also help me to understand the sport better."

The other two goals that were not technical in nature described soft skills as their goals, as indicated by S02 and S19:
- S02: "I want to learn on how to improve on time management by doing a survey. I wish to do so as I would want to implement an effective method of time management for my daily life as I would usually procrastinate on work and rush things last minute, stressing me out."
- S19: "Presentation skills for analytics. I feel that to be able to show and explain data through words and body language is a good skill for this module, and that I should improve in my current abilities as I have stuttering problems and get flustered during speeches."

At the end of the assignment, the participants are asked to complete a reflection survey to gather their thoughts on the whole learning process.  Only 24 participants' responses are analyzed because two of the 26 participants who completed both the assignment and survey submitted identical survey responses.  Hence, both these responses are discarded.

The survey question *Which part of this assignment makes sense to you?* is set as open-ended for participants to share their feedback on this independent learning journey.  Based on past experience of doing similar studies, it was expected that if participants resisted such assignment, their open-ended responses would indicate their objections towards the transfer of responsibility of learning from the teacher to the learners themselves, or that negative sentiments would surface.  As the survey carried no additional marks to their overall score, the participants are encouraged to express their views honestly.

There was no evidence of negative sentiment from participants' open-ended responses.  In fact, approximately 30% of the participants who completed the reflection survey also described parts of SDL in the assignment that made sense to them.  Such responses are shared by participants S13, S18, S28 and S34:

- S13: "Actually I have never try to do this kind of assignment before so it is like a whole new experience. At this point of time, reading through some of the goals i write many weeks ago is kind of amazing."
- S18: "When you're marked based on goals which we set on our own. This really shows whether someone took the initiative and heart to put and learn something when given the opportunity to."
- S28: "(…) The setting of goal showed me how I am achieving it like what mistake did I did wrong and what I am success."
- S34: "Setting milestones for myself and grading myself."

As a large majority of the goals were directed at learning technical skills, it is not surprising that what made sense to the participants are also largely associated with the attainment of technical skills.  Response by S38 gives one such example: "Everything made sense. From the visualisation to pre attentive features and collecting of data."

As a summary, three-quarters of the participants in this study embraced an increase in autonomy in learning, and approximately 30% of these participants also seem to be in agreement that an increased in autonomy is associated with positive perceptions of SDL.

The following section presents the findings on how participants made sense of self-directed learning.  First, *Figure 2* summarizes the quantitative data aggregated using simple descriptive statistics and establishes two clusters of participants in responding to the tasks with increased autonomy.  Then, *Figure 3* maps the analysis from the quantitative data to the coded qualitative data and offers detailed discussions on how the participants internalized SDL through the assignment.

*What are the differences in how learners make sense of SDL in a module designed using Ryoo's (2011) minimalist view on autonomy?*

*Table 2: Numerical summaries for goal setting and monitoring of learning (e-journal submissions)*

| | No. of e-journal submissions | Lapses between goal-setting and first e-journal submission (in days) | Median days between e-journal submissions |
|---|---|---|---|
| Total | 128 | - | - |
| Minimum | 1 | 1 | 0 |
| Median | 4 | 55 | 3 |
| Maximum | 10 | 70 | 68 |

Table 2 gives a general overview on number of e-journal submissions, the lapses in days between participants' goal setting and subsequently monitoring of their learning through e-

journal submissions. The participants' first e-journals are submitted generally about 50-70 days after the goals are set.  This may imply that perhaps the participants believed such assignment should not require prolonged effort and sustained interest and would not warrant such demands too.  In addition, participants seemed comfortable and confident with short-term goals as seen in the long lapses between goal setting and first e-journal submissions.  Furthermore, the participants commonly submitted one to 10 e-journals and the aggregated number of days that lapsed between two e-journal submissions is approximately 3 days.  It was of interest to analyze deeper how participants spent their time doing this assignment.

Figure 2 shows further analysis on the relationship between three variables - Lapses in days between goal setting and the first e-journal submission (x-axis), median days between e-journal submissions (y-axis) and the number of e-journal submissions (size of the bubble). At least two clusters seem to appear - one cluster indicated by bubbles at the bottom right hand corner in, coded as Delayed-Start and another on the left side in Figure 2 coded as Prompt-Start.

Delayed-Start cluster indicates participants who submitted their first e-journals approximately 50-70 days after the goals are set.  About 60% of the participants fall in this cluster.  There are small bubbles and bigger bubbles in this cluster.  Smaller bubbles indicate fewer e-journal submissions.  Most of the smallest bubbles like the one indicated by 'A', are in this cluster.  This means the participants' first e-journals are also their only e-journals. Participants from this group likely forgot about the deadline of this assignments, sought quick closure to complete the assignment when the deadline is near, and/or could likely have over- estimated their capabilities in completing the assignment as desired.  A few bubbles in the Delayed-Start cluster are also big in size, some resting very near the horizontal axis.  This means the participants submitted many e-journals within a very short period of time.  It may be possible that participants did their reflections in another platform and transferred their reflections over to LMS at the end of the assignment period. Another possible explanation hints at participants tried to gain more credit by compensating for their lack of regular effort with high number of submissions at the end.  Generally, Delayed-Start cluster submissions imply lack of evidence in consistently monitoring their learning effort.
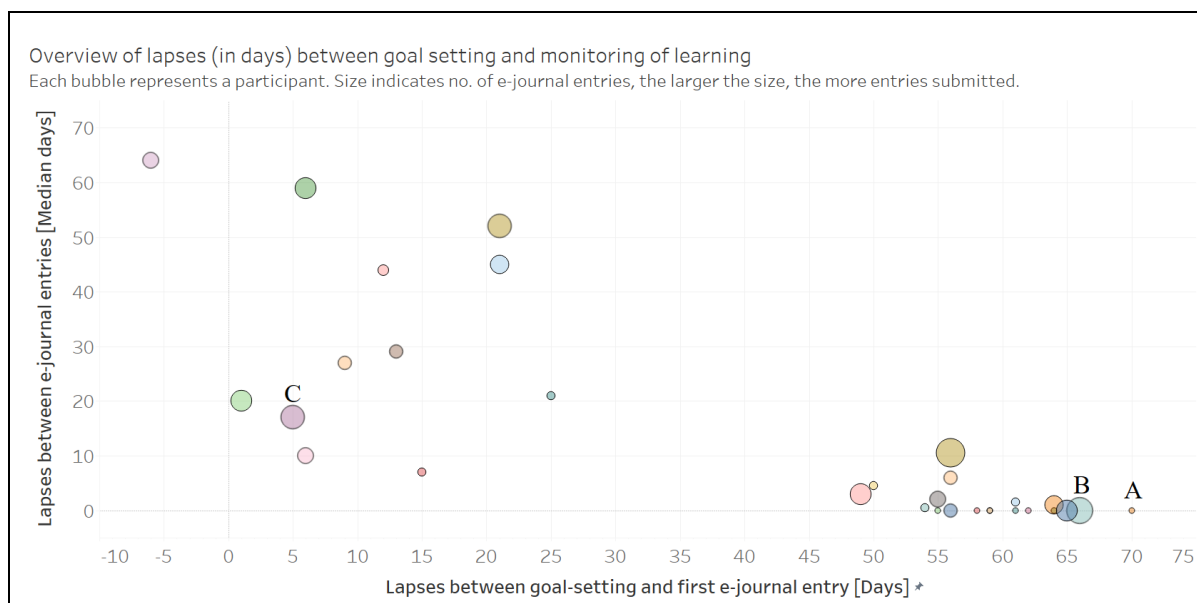
*Figure 2. A comparison between the lapses in number of days between goal setting and first e-journal submission (x-axis), median number of days between journal submissions (y-axis) as well as the number of submissions that each participant submitted (size of the bubble - the bigger the size, the higher number of submissions).*

*Prompt-Start* cluster indicates participants who submitted their first e-journals within a month of goal setting.  Unlike *Delayed-Start* cluster, the median number of days within *Prompt-Start* e-journal submissions were also more dispersed.  Bubbles towards to the top left-hand corner may indicate longer breaks in between e-journal submissions.  Bubbles bigger in size towards bottom left-hand corner are the most desired, like those indicated by 'C'.  These bubbles showed evidence of more frequent e-journal submissions, which could suggest more regular reflection of their learning.  The aggregated median number of days between e-journal submissions coincides with two to four major learning units in this module.  This suggests that some participants in the *Prompt-Start* cluster can set goals and monitor their own learning consistently.

Subsequently, participants' qualitative responses from the e-journal submissions and final survey are scrutinized for evidence on how they have monitored and evaluated their own learning, as well as how they have internalized the concept of SDL through this assignment. *Figure 3* uses a parallel plot to map the quantitative data analyzed (verticals D to G) to the coded qualitative data (verticals H to J).
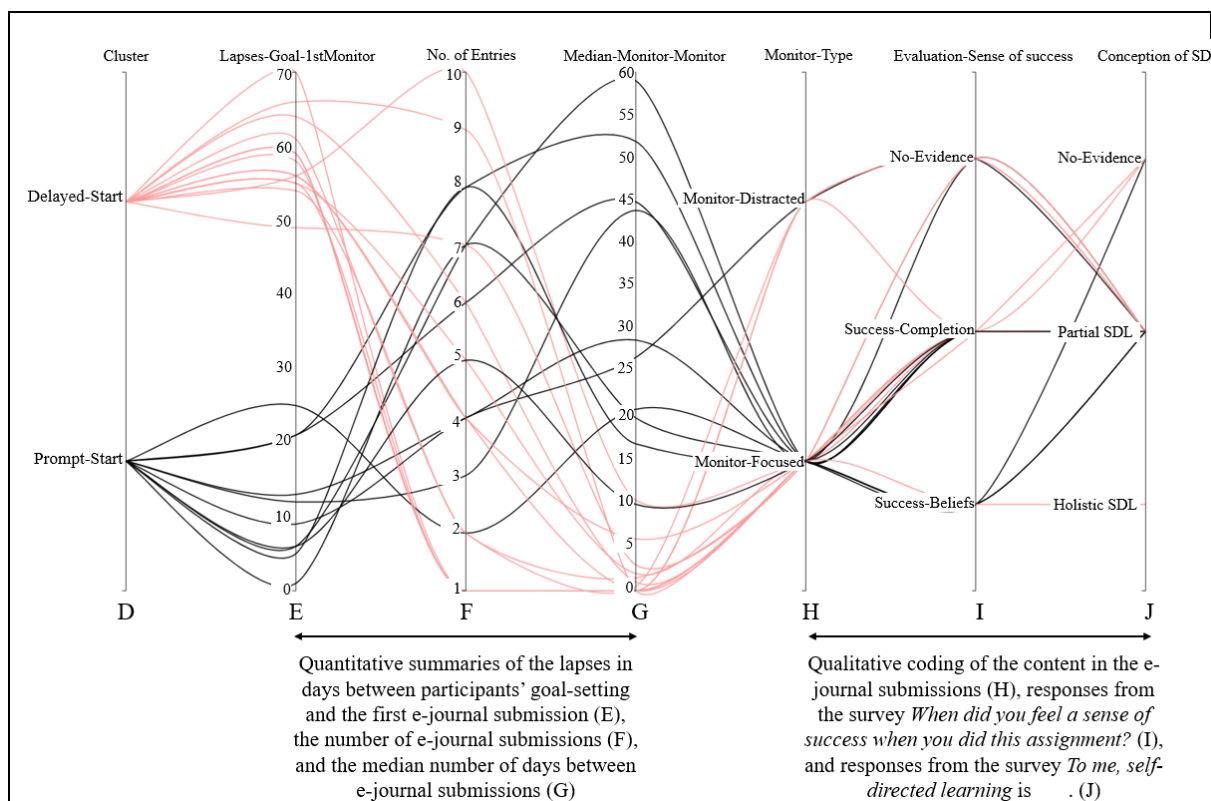
*Figure 3. Each line traces the characteristics of one participant's monitoring of own learning (verticals F to H), evaluation of own learning (vertical I), and ultimately his or her conception of SDL as experienced through this assignment (vertical J). The lighter line represents Delayed-Start participants who submitted their first e-journal about 50-70 days since goal setting, whereas the darker line represents Prompt-Start participants who submitted their first e-journal within a month of goal-setting (verticals D and E).*

*Monitoring own learning (Figure 3 verticals F to H).* Generally, *Prompt-Start* participants submitted a slightly smaller number of e-journals compared to *Delayed-Start* participants. Almost all the e-journal submissions are focused on achieving the goals set. These e-journals are coded as *Monitor-Focused* where the goals set at the beginning are evident and learning done are directed towards achieving the goals. There is also evidence that the participants know how to self-evaluate their own learning by awarding marks to themselves when a milestone set has been achieved. A summarized example of *Monitor-Focused* submissions is given by S33 during the assignment period: "I watch 2 YouTube video on how to create animated visuals in tableau, which I am supposed to complete it by week 6 as part of my milestone 1 (…) I have shown [tutor] the animation hence milestones 1 and 2 are completed. Marks awarded to myself: 3 (…) i have completed my milestone 3 which was to create a animated visual for my group project (…) Marked awarded to myself : 2 marks. i have completed all my milestones."

On the other hand, e-journals coded as *Monitor-Distracted* do not show evidence of the goals set as the center of focus in their learning effort. The common reason given is the lack of proper time management - where they admitted that they forgot about the assignment,

or that they decided to work on assignment which they place higher values on.  An example of an e-journal submission that was coded as *Monitor-Distracted* is given by S38: "This is my first journal entry for the Assignment Milestone which was supposed to keep track of my progress throughout the making of the dear data (…) The whole process of dear data should have been done by 31 May 2019(…) i became a little busy and occupied by my other modules and completely lost track of time (…) Today, 3/7/2019, i have only started to start sketching my raw data on a postcard to be given to [tutor] just in time for my submission."

Although most of the e-journals submissions are focused on the goals set, those submitted by *Prompt-Start* participants show median days between e-journal submissions to be in the range of 10 to slightly below 60 days.  As such there could be evidence of a more distributed practice or reflection on their learning and working towards the goals.  As a comparison, *Delayed-Start* participants generally has a much lower median days between e-journal submissions, many show median to be zero day.  This means there was mass submissions within a very short period of time, or that the participants only submitted one e-journal.  This point was also shown in *Figure 2* and discussed in the previous section.

*Evaluating own learning (Figure 3 vertical I).* In terms of evaluating their own learning, some 58% or 14 out of 24 participants reported that they felt a sense of success when they have completed the assignment.  The completion tasks, especially those that they are able to compare against an external standard, give the participants a sense of achievement regardless of whether the participants started on their goals promptly or not, and how frequently they submitted e-journals to monitor their progress.  Participants S28 and S32 share their reflections on when they felt a sense of success when doing this assignment:
- S28: "When I complete the drawing and by looking at it, it feels like is almost the same as the dear data website sample."
- S32: "I feel a sense of accomplishment when I finish my storyboard and see the things I learn are put into use."

However, a few participants did not provide convincing evidence that they felt a sense of success.  When probed further, one of the reasons given is that they did not manage to achieve the goals set.  Responses by S35 and S27 provide such examples:
- S35: "I feel that the amount of effort I had put in could have been more as I did not manage to achieve the goals I had set."
- S27: "I was not able to use the data to find out the difference in what makes a good climber and what makes a bad climber."

Some other participants who reported a sense of success are also coded as *No-Evidence* because the achievement is focused on attaining the outcome of an unrelated goal.  There is no evidence that the goal set in the beginning has been achieved or that this achievement gave them a sense of accomplishment.  An example is given by S02 who had set goals on improving his time management but the sense of success he reported is when he was able to prepare the data set properly for his project.  There is no indication whether his goal in managing time better is attained.  These participants whose responses on sense of achievement are coded as *No-Evidence* are largely from the *Delay-Start* cluster.  Perhaps

because they had delayed the start of working on their goals, there is no ample opportunity for them to gather some initial evidence to assess the feasibility of the goals that they set, re-evaluate the resources needed to achieve those goals, or even to regulate their thinking on whether their effort is directed appropriately towards achieving the goals set.

In comparison, a few participants from the *Prompt-Start* cluster reported a sense of success not only in the completion of the tasks but also in how the completion of the tasks changed their prior beliefs about their own capabilities and challenges they assumed the tasks held. Examples are given by S18 and S33:

- S18: "I've managed to [l]earn something out of class, by doing my own respective research, which made me proud when I've managed to achieve what i set out to do from the beginning."
- S33: "I feel a sense of achievement, initially I felt that creating animated visual is a challenging process, I doubted myself. However, I am able to create one animated visual."

*Conception of SDL (Figure 3 vertical J).* In general, almost all the participants' conception of SDL are coded as *No-Evidence* and *Partial-SDL*. These responses originated both from the *Delay-Start* and *Prompt-Start* cluster. In other words, there seem to be little evidence from this study that the promptness in getting started on achieving a goal, the consistency in monitoring own learning and the confidence to evaluate own learning are strong differentiators in how participants conceptualized SDL. Two possible explanations could be offered for this observation. The first explanation may be linked to the limitation of collecting data from a typed-written survey in which participants may not be predisposed to elaborate on their conceptions of SDL. The second explanation may be gleaned from tracing the characteristics of monitoring and evaluating own learning of the response coded as *Holistic-SDL*. There seems to be some evidence to suggest that short-term intensive and focused practice, rather than longer term prolonged and sustained effort may likely lead to a more holistic view of SDL. Conception of SDL coded as *Holistic-SDL* emphasizes an end-target and stresses on the self in the monitoring and evaluating of own learning. Response by S18 is an example: "To me, self-directed learning is taking the initiative to learn and enhance your knowledge at will. It is also when you need to have a goal set out in mind, and know what you want/need to do. Then focus on what you've set out to do, and follow it with a end goal in mind. Only then, upon accomplishment on what you've set out to do during SDL, will you feel a sense of achievement."

Figure 4 traces S18's evidence from his qualitative and quantitative responses on monitoring and evaluation of learning, and then a mapping to his conception of SDL. It is noteworthy to observe that S18 whose SDL conception coded as *Holistic-SDL* is from the *Delayed-Start* cluster. S18 submitted five e-journals containing evidence of focused learning effort directed at achieving the goal set. The median days between e-journal submissions is two. S18 reported a sense of success when he could independently extend his learning beyond the class by setting and achieving goals.
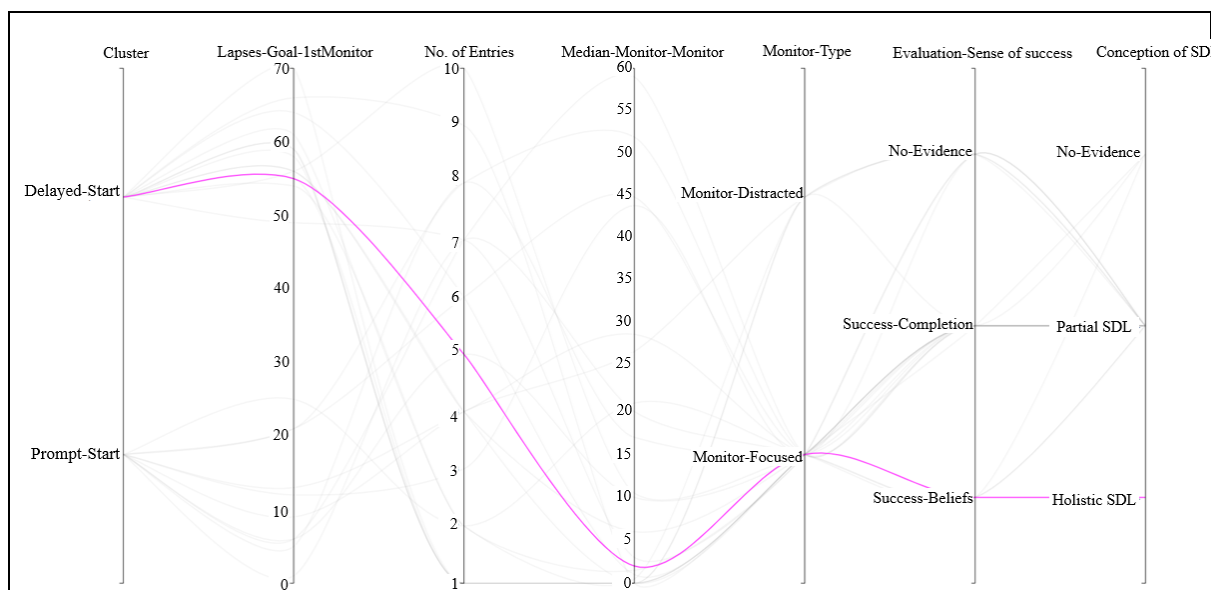
*Figure 4*. S18's monitoring and evaluation of learning and the mapping to his conception of SDL.

The data collected from S18 seem to suggest that focused and intensive effort regardless of how promptly they start working on their goals could be associated with a more holistic conception of SDL.   Perhaps the gratification derived from achieving a short-term goal serves as validation to the capability of independent learning and could likely spur subsequent SDL.  Hence, the pre-conceived notion that consistent effort over a pro-longed period of time needs for SDL to be developed should be re-examined as short-term intensive and focused effort could also be linked to a more holistic conception of SDL, as portrayed by S18.

Approximately 80% of the participants provided evidence that they could describe some important components of SDL such as the emphasis of self in taking the initiative, with or without the help of others, in goal setting, monitoring, and evaluating their own learning. But at the same time, such responses also lack the mention of self as just as an important source of validation of their own learning.  It seems that a majority of the participants have not yet internalized that evaluation, especially self-evaluation, is an important component in SDL to validate those goals have been achieved through intentional and purposeful learning. As such, these responses are coded as *Partial-SDL*.  Some examples are given by the S04, S29 and S33:

- S04: "Taking the initiate and motiving yourself to do it"
- S29: "[B]eing able to be independent and learn on your own without the need of a lecturer to guide you."
- S33: "Is having a good time management and having the right mindset. Step by step process to track on my learning."

Three out of 24 participants responded in a general way of what SDL meant to them:
- S13: "[A] new way of learning"
- S28: "Important and it can be difficult at the start"

- S30: " [A] test for self-discipline"

Since their responses on SDL were rather general and did not make references to the important components of goal setting, monitoring, and evaluating learning, these are coded as *No-Evidence*.  The omnipresence of an end-goal is important because it directs learning effort purposefully, serves as a validation for learning and subsequently when end-goal is attained, it releases a sense of achievement that may motivate or reinforce a desired follow-up learning behavior.  All these are contained in self and can be activated with or without the need for direction exerted by an external force.  As such, participant may then get a sense that SDL is a skill regardless of the type of learning to be done, if the process components of goal setting, monitoring and evaluation are in place.

**Conclusion**

This study offers three important insights into incorporating autonomy for self-directed learning.  First is in the use of simple visualizations to inform teachers about their learners SDL characteristics.  This study uses a scatterplot in *Figure 2* as well as parallel plots in *Figure 3* and *Figure 4* to trace each participant's quantitative and qualitative responses and mapped it to their conceptions of SDL.  Both these visualizations can be generated effortlessly from readily available commercial visualization software like Tableau, Power BI and KNIME.  Second, a minimalist approach to autonomy incorporated in tasks could encourage learners to develop substantial conception of SDL.  The most important element is then to encourage participants to always keep the end-goal in mind, and to direct effort purposefully in achieving the goal.  As a side note on this point, the role of technology is important in both supporting the increase in autonomy by documenting and managing evidence of learning properly to avoid any conflict in the evaluation phase.  This study choses a LMS system for such purpose.  Lastly, the findings in this study seem to suggest that learning tasks that incorporate autonomy for self-directed learning could be designed as short-term tasks that require intensive and focused effort rather than one that is requires prolonged effort and sustained interest.

**About the Author**

**Hui Teng Chia, Ph.D.** is a Teaching and Learning Mentor at Singapore Polytechnic in the Analytics section of the School of Mathematics and Science.  Besides teaching statistics and analytics modules at the polytechnic, she mentors new academic staff and guides peers in action research and reflective practices.  Her research interests include how people make sense of data and visualizations in teaching and learning.  Email: chia_hui_teng@sp.edu.sg

## References

Balser, T. (2018). *Beyond content: Teaching in the post-modern classroom*.  Keynote presentation at the 2018 Excellence in Education and Teaching Convention, Singapore Polytechnic, Singapore.

Bartholomew, S. R., Reeve, E., Veon, R., Goodridge, W., Lee, V. & Nadelson, L. (2017).  Relationships between access to mobile devices, student self-directed learning, and achievement.  *Journal of Technology Education*, *29*(1), 2-24.

Douglass, C. & Morris, S. R. (2014).  Student perspectives on self-directed learning.  *Journal of the Scholarship of Teaching and Learning*, *14*(1), 13-25.

Gibbons, M. (2002).  *The self-directed learning handbook: Challenging adolescent students to excel*. San Francisco: Jossey-Bass.

Knowles, M. S. (1973).  *The Adult Learner: A neglected species*.  Retrieved from: https://files.eric.ed.gov/fulltext/ED084368.pdf

Knowles, M. S. (1975).  *Self-directed learning: A guide for learners and teachers*. New York: The Adult Education Company.

Knowles, M. S. (1980).  *The modern practice of adult education: From pedagogy to andragogy, revised and updated*.  Retrieved from: https://pdfs.semanticscholar.org/8948/296248bbf58415cbd21b36a3e4b37b9c08b1.pdf

Ministry of Education, Singapore (2018). *21st Century Competencies*.  Retrieved from: https://www.moe.gov.sg/education/education-system/21st-century-competencies

Peine, A., Kabino, K. & Spreckelsen, C. (2016).  Self-directed learning can outperform direct instruction in the course of a modern German medical curriculum - results of a mixed methods trial.  *BMC Medical Education*, 16:158.

Piskurich, G. M. (1993). *Self-directed learning: A practical guide to design, development, and implementation*.  San Francisco: Jossey-Bass.

Van der Walt, J. L. (2019). *The Term "Self-Directed Learning"--Back to Knowles, or Another Way to Forge Ahead?* [Abstract]. *Journal of Research on Christian Education*, *28*(1), 1-20.

Russo, A. (2020, October 20).   *Recession and Automation Changes Our Future of Work, But There are Jobs Coming, Report Says*.  World Economic Forum. https://www.weforum.org/press/2020/10/recession-and-automation-changes-our-future-of-work-but-there-are-jobs-coming-report-says-52c5162fce/

Ryoo, J. S. C. (2011).  An analysis of practical autonomy in education: The case of Korean self-directed learning policy.  *KEDI Journal of Educational Policy*, *8*(2), 345-363.

# SPANISH VOCABULARY ACQUISITION BY ENGLISH SPEAKERS USING SPACED-REPETITION ALGORITHMS AND MNEMONICS

**Jon-Erik G. Storm**
**Nipomo High School**

**Abstract** Two popular techniques for memorizing vocabulary, both in general and in foreign-language learning, are the use of spaced-repetition software and mnemonic phrases. The author gave three classes of first-year Spanish students in high school 20 new high-frequency vocabulary words each Monday, provided 15 minutes of time each day to use spaced-repetition software to memorize them, and quizzed the students on these words at the end of each week. Performance on orthographically similar words was 21% stronger than words that were not orthographically similar prior to the mnemonic intervention described here. The experimental group's performance on words that were not orthographically similar showed statistically significant improvement after the intervention. The control group's performance did not. This suggests two conclusions (1) that the simple addition of mnemonics to words that are not orthographically similar increases vocabulary retention and (2) spaced-repetition algorithms do not adequately account for differences between words of difficulty varying on the basis of orthographic similarity.

**Keywords:** teacher action research, Spanish, mnemonics, second language acquisition, spaced repetition

## Introduction

Two popular strategies for memorizing vocabulary, both in general and in foreign-language learning, are the use of spaced-repetition software (SRS) (Ono, 2017) and mnemonic devices (Paivio & Desrochers, 1981). Appropriate use of SRS should result in learners seeing words that are more difficult for them more often. While the literature suggests that employing a variety of language-learning strategies results in the best outcomes (Gholami, Abdorrahimzadeh, & Behjat, 2014), given a finite amount of time and other practical constraints, finding optimal mixes of language-learning strategies can make optimal use of time and other resources. In this paper, the author finds that SRS usage can be improved on by short interventions with mnemonics on words that are not orthographically similar in L1 and L2.

**Literature Review**

The second-language acquisition (2LA) literature has individually studied both SRS and mnemonics quite well. Before now, however, there is no known study of the combination of these two strategies in the context of 2LA. Outside of 2LA, one study showed that a combination of both enabled effective memorization of strong passwords (Blocki, Komanduri, Cranor, & Datta, 2014). Seibert Hanson and Brown suggested as a future research direction to analyze the combination of SRS and mnemonics (Seibert Hanson & Brown, 2019). This paper seeks to advance that research.

SRS uses algorithms based on brain research into the mechanics of human memory. Some examples of software that uses SRS algorithms assessed in research include Anki, Memrise, and SuperMemo. Numerous studies have shown SRS to be effective in acquiring L2 vocabulary and to foster a growth mindset in learners and to stimulate learner motivation (Seibert Hanson & Brown, 2019; Ono, 2017). SRS uses flashcards but spaces out reviews based on the learner's subjective estimate of their performance. If the correct answer immediately pops into the student's mind, the word gets marked "easy"; if the correct answer occurs to the student in a second or two, the student marks the word "good"; if the student answers correctly after some time, it gets marked "hard." If the student does not get the answer correct, the student selects "again." The algorithm spaces out reviews of the cards based on these inputs with incorrect answers being shown again in the same session and "easy" words being delayed for review for an increasingly long period of time. SRS research suggests that this creates ideal results for vocabulary retention (Seibert Hanson & Brown, 2019). Despite these results and the algorithmic optimization used by SRS, Ono 2017 still found that word length and prior language experience affected the results. Similarly, this study shows the degree of orthographic similarity creates a disparity of performance when using SRS.

For over a century, language instructors have been using mnemonics to improve retention of L2 vocabulary. Some techniques for creating mnemonics include using acoustic, orthographic, or semantic "links," or, failing that, a picture (Paivio & Desrochers, 1981). For example, for Spanish *perro* (*i.e.*, dog pronounced very roughly like "pear-oh"), we might draw a picture of a pear-shaped dog. The pear-shape is an acoustic link. Research shows that the use of mnemonics, whether created by the learner or the instructor, significantly improve recall (Paivio & Desrochers, 1981).

Vocabulary is perhaps the most important area of language acquisition, and one that is perhaps the most frustrating for learners (Meara, 1980). By emphasizing high-frequency vocabulary, instructors and learners can leverage Zipf's law, which stands for the proposition that most of speech and writing is comprised of a small set of high-frequency lemmata (*i.e.,* groups of word forms that differ only by grammatical prefix or suffix) (Nation & Waring, 1997). The result is that L2 vocabulary acquisition can optimized by learning only the most frequently used lemmata in a language and the memorization of those lemmata can be optimized using SRS.

Educators can present students with numerous language-learning strategies, but thereby risk losing focus, confusing students, and depleting time. By finding a mix of strategies that effectively uses time, educators can make the best use of time. The purpose of this study is to investigate the mix of SRS and mnemonics in 2LA and to make an original contribution to 2LA pedagogy through a rigorous statistical analysis of the results.

**Methodology**

*Participants.* The author gave 68 first year high-school Spanish students in the author's three class periods 20 new high-frequency vocabulary words each week, starting with the most frequent at the beginning of the year and progressing towards a goal of completing the top 500 most common Spanish lemmata by the end of the year, which accounts for approximately 50% of the words used in speech and writing. The author quizzed the students on 15 of the 20 new words, 5 words from the prior week, and 5 randomly chosen from weeks prior. Participants were approximately 3% African American, 12% Asian-American, 38% Latino, and 47% Caucasian. Participants were 46% female and 54% male. Ages range from 14 to 17. All participants had completed one semester of high school Spanish I. Each class period, students are given approximately 15 minutes at the beginning of class to study the words with SRS.

*Design.* The author used an independent but not randomly assigned two group design. The author designated two classes as the experimental group and one as the control. The experimental group was presented with the mnemonic intervention and the control group was not. After reviewing the data, the author finds that there is virtually no difference between the demographics or academic achievement of the two groups. All other instruction, content, lesson plans, and other variables were held constant between the groups.

*The Intervention.* In preparation for the second quiz, the experimental group was presented with mnemonic sentences (*e.g.,* "*Suceder* / To Happen: He made it *happen* because he was a "succeed-er") for the non-orthographically similar words only. The control group was not. The mnemonics were repeated twice each class period Monday through Friday, taking about 2 minutes per day. Otherwise, the preparation for both groups was identical. The author created the mnemonics. The author presented the mnemonics to the entire class at the conclusion of the time allocated for spaced repetition software use. The mnemonics were written on the back of flashcards and presented in the following manner: (1) the word was read aloud in Spanish; (2) the word's definition in English was read aloud; (3) followed by a mnemonic story or pun that links the word. For example, "*Fuera*. Outside. She wanted him to go FAR "FUER"-A-way, so he went OUTSIDE." The stories do not need to be sophisticated or clever. Simply linking the two words, even absurdly, is sufficient. After the author read each card twice, 5-10 random students were cold called to check if they remembered the mnemonic.

*Quiz Instrument.* The quizzes tested 15 new words each week and 5 words chosen from the previous weeks' sets. The further 5 words from past weeks chosen at random are not

analyzed here. Students simply translate the words from English to Spanish and vice-versa. A correct answer may include mild typos if it does not collide with another word. On that basis, the answers are either scored correct or incorrect. There is an equal number of "active" translations from English to Spanish and "passive" translations from Spanish to English. All classes took the same quizzes.

To analyze the quizzes, the author sorted words into two study categories: (1) orthographically similar words (*e.g.,* possibility and *posibilidad*); (2) words that are not orthographically similar (*e.g.,* weight and *el peso*).

**Results**

Because the data was not clearly normally distributed and the data were paired across the quizzes, the author employs the Wilcoxon Signed Rank Test, as implemented by SPSS software throughout. The small sample size is a result of the smaller student body at the author's school, but it is still sufficient in size to test the hypotheses using the Wilcoxon Signed Rank Test.

As a null hypothesis, we assume that there should be no meaningful change in the control group on non-orthographically similar words between Quiz 1 and Quiz 2.

*Table 3: Related-Samples Wilcoxon Signed Rank Test Summary Between Quiz 1 and Quiz 2 For the Control Group (Non-Orthographically Similar Words)*

| | |
|---|---|
| Total N (students) | 19 |
| Test Statistic | 102.000 |
| Standard Error | 21.119 |
| Standardized Test Statistic | 1.207 |
| Asymptotic Sig. (2-sided test) | .227 |

Based on the data in Table 1, we retain the null hypothesis because the asymptotic significance exceeds 0.05. Using the same null hypothesis for the experimental group and an alternative hypothesis that the intervention should result in a difference, we reject the null hypothesis with an asymptotic significance of less than 0.05 in Table 2.

*Table 4: Related-Samples Wilcoxon Signed Rank Test Summary Between Quiz 1 and Quiz 2 For the Experimental Group (Non-Orthographically Similar Words)*

| | |
|---|---|
| Total N (students) | 49 |
| Test Statistic | 685.000 |
| Standard Error | 85.653 |

| | |
|---|---|
| Standardized Test Statistic | 2.218 |
| Asymptotic Sig. (2-sided test) | .027 |

As a further control, we can see whether there was a meaningful difference between words that were orthographically similar. In both experimental groups (Table 3), control (Table 4), and in the aggregate (Table 5), we retain the null hypothesis of no meaningful differences between the quizzes.

*Table 5: Related-Samples Wilcoxon Signed Rank Test Summary Between Quiz 1 and Quiz 2 For the Experimental Group (Orthographically Similar Words)*

| | |
|---|---|
| Total N (students) | 49 |
| Test Statistic | 259.000 |
| Standard Error | 43.859 |
| Standardized Test Statistic | 1.277 |
| Asymptotic Sig. (2-sided test) | .202 |

*Table 6: Related-Samples Wilcoxon Signed Rank Test Summary Between Quiz 1 and Quiz 2 For the Control Group (Orthographically Similar Words)*

| Related-Samples Wilcoxon Signed Rank Test Summary | |
|---|---|
| Total N (students) | 19 |
| Test Statistic | 35.000 |
| Standard Error | 9.747 |
| Standardized Test Statistic | .769 |
| Asymptotic Sig. (2-sided test) | .442 |

*Table 7: Related-Samples Wilcoxon Signed Rank Test Summary Between Quiz 1 and Quiz 2 For All Students (Orthographically Similar Words)*

| Related-Samples Wilcoxon Signed Rank Test Summary | |
|---|---|
| Total N (students) | 68 |
| Test Statistic | 1094.000 |
| Standard Error | 135.444 |

| Standardized Test Statistic | 1.322 |
|---|---|
| Asymptotic Sig. (2-sided test) | .186 |

**Discussion**

The only statistically significant result is in the performance of the experimental group on words that are not orthographically similar. This is the group that received the simple, 2-minute mnemonic intervention and on the group of words that is the target of the intervention. Because the intervention narrows a previously existing performance gap between orthographically similar words on the one hand, and those that were not orthographically similar on the hand, the data also suggest that SRS alone may not sufficiently practice non-orthographically similar words. It seems possible that the orthographic similarity itself serves as a "built-in" mnemonic. It is unclear why either SRS does not self-adjust for this or why learner input does not reflect this sufficiently for SRS to self-adjust for it.

**Implications**

This is a small data set from a small group of students too small to be representative of all learners, even of all high-school Spanish-learners. The sample size is sufficient to test the hypotheses using appropriate statistical methods, but not sufficient to cross-tabulate on the basis of demographics, gender, academic achievement, or other factors. The demographic mix in this study is quite different from the country as a whole. It is also quite different from the demographic mix of Spanish learners as a whole and only includes Latino and Caucasian students in significant numbers. Despite these limitations, the result is clear: augmenting SRS with a brief mnemonic intervention of only the non-orthographically similar words improves retention.

Future research should, of course, use larger sample sizes and, if possible, random selectees with a more representative demographic mix. In addition, specifications for what makes mnemonics most optimal in this context and some method of measuring orthographic similarity should be developed. Other questions include whether improving student training with SRS to give more accurate input, if indeed their input is not correct, will result in the algorithm self-adjusting to compensate for the "built-in mnemonic" of similar spelling.

**Conclusion**

Both SRS and mnemonic devices are popular language-learning strategies. While the use of a variety of language-learning strategies results in the very best outcomes, time and other practical constraints in the classroom call out for time-optimizing approaches to the use of these strategies. This paper has shown that short interventions with mnemonics on words that are not orthographically similar in L1 and L2 can improve outcomes compared with SRS usage alone.

**About the Author**

**Jon-Erik G. Storm, J.D.** teaches mathematics and Spanish in a rural school in Nipomo, California. He holds a doctorate in environmental and energy law and policy. His recent research focuses on technology in teaching. Email: jon-erik.storm@lmusd.org

## References

Blocki, J., Komanduri, S., Cranor, L., & Datta, A. (2014). *Speced Repitition and Mnemonics Enable Recall of Multiple Strong Passwords.*

Gholami, M., Abdorrahimzadeh, S., & Behjat, F. (2014). L2 Literacy Proficiency And Language Learning Strategies. *Modern Journal of Language Teaching Methods, 4*, 77-90.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Teaching, 13*, 221-246.

Nation, P., & Waring, R. (1997). Vocabulary Size, Text Coverage And Word Lists. In N. Schmitt, & M. McCarthy, *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6-19). Cambridge University Press.

Ono, T. (2017). Vocabulary Learning Through Computer Assisted Language Learning. *Hitotsubashi Journal of Arts And Sciences, 58*(1), 67-72.

Paivio, A., & Desrochers, A. (1981). Mnemonic Techniques in Second-Language Learning. *Journal of Educational Psychology*, 780795.

Seibert Hanson, A. E., & Brown, C. M. (2019). Enhancing L2 learning through a mobile assisted spaced-repetition tool: an effective but bitter pill? *Computer Assisted Language Learning*.

Zuckermann, G. (2011). Mnemonics in Second Language Acquisition. *Word Ways, 44*(4).

# COLLABORATIVE TESTING IMPACTS STUDENT ACHIEVEMENT AND TEST ANXIETY FOR ADVANCED PLACEMENT ENVIRONMENTAL SCIENCE STUDENTS

**Helen Bremert**
**Sydney University**


**S.B. Boesdorfer**
**Illinois State University**

**Abstract** Assessment is a concern in high schools because summative assessments can impact a student's potential. This researcher studied how collaborative assessments increased student achievement and reduced test-taking anxiety after observing students struggling when undertaking high stakes testing. I draw on previous research into collaboration to demonstrate the benefits of collaborative assessment. The study occurred in a suburban high school AP Environmental science class. Students undertook six assessments, three individual and three collaborative high stakes tests in heterogeneous groups. A repeated-measures ANOVA test and paired t-tests were conducted to determine the differences between the testing methods. An analysis of the findings indicates there is a significant difference between students undertaking individual and collaborative tests. In the post-test Qualtrics survey, many students showed an increased understanding of the content and reduced their test-taking anxiety. The paper argues that collaborative assessment has many benefits to students; it increases students' achievement and understanding of the concepts as they utilize reasoning and argumentation to defend their answers. Additionally, the negative impacts associated with collaborative testing can be easily addressed.

## Introduction

Education in America has changed with the recognition that students must be productive citizens in a rapidly changing world (Achieve, 2010). Scientists and other professionals worldwide require employees to collaborate, think critically, and problem-solve effectively

(Ngotngamwong, 2014). To this end, the *Next Generation Science Standards* (NGSS) has a vision of ensuring that upon graduation, students will have the necessary skills to be practical and rational thinkers (NGSS Lead States, 2013). To achieve this, the NGSS recommends inquiry, collaborative, and evidence-based instruction across a wide range of science fields (Nairman & Chrispeels, 2016; NGSS Lead States, 2013).

Collaborative learning is pedagogical method that promotes an active classroom learning environment where the students form pairs or groups to accomplish tasks (Meseke, Nafziger & Meseke, 2010). This instruction method's advantages include, increased conceptual understanding, retention, problem-solving, and critical thinking skills (Gilley & Clarkston, 2014; Pandey & Kapitanoff, 2011; Siegal, Roberts, Freyermuth, Witzig & Izci, 2015). Furthermore, collaboration promotes heightened intrinsic motivation, interpersonal skills, and students' ability to engage in evidence-based argumentation (Guiliodori, Lujan & DiCarlo, 2008; Meseke et al., 2010; Zipp, 2007), which are all necessary skills for the current workforce and supported by NGSS.

During high school, many students engage in collaborative tasks and formative assessments; however, when students undertake summative assessments, they are generally taken individually (Siegal et al., 2015). Quizzes and examinations occur because they can assess students quickly and over many learning units (Rao, Collins & DiCarlo, 2002). Additionally, this method enables the educator to determine students' academic strengths and weaknesses and holds schools accountable. Nonetheless, individual tests have several disadvantages, including lowering intrinsic motivation, using only information-recall type questions, underperformance due to outside factors, and increased test anxiety (Breedlove, Burkett & Winfield, 2004). An alternative testing method used to combat issues with individual testing is collaborative assessments. One such testing issue is the reduction test taking anxiety (Haberyan & Barnett, 2010; Ngongamwong, 2014; Siegal et al., 2015) due to students being able to "have the emotional and intellectual support" (Rao, Collins & DiCarlo, 2002, p. 38) of their peers. Moreover, group testing may improve exam performance (Gilley & Clarkston, 2014; Vogler & Robinson, 2016) and promote positive student attitudes (Haberyan & Barnett, 2010).  In a group testing environment, the students must discuss questions and answers, thereby filling in knowledge gaps, leading to greater understanding and greater retention of the material (Kapitanoff, 2009; Vogler & Robinson, 2016). Much of the research on the effects of collaborative testing (Breedlove et al., 2004; Gilley & Clarkston, 2014; Meseke et al., 2010; Rao et al., 2002; Siegal et al., 2015; Vogler & Robinson, 2016) has focused on undergraduate students. Despite the positive outcomes of collaborative testing at the collegiate level, there is little data on how high school students' summative performance could improve through collaborative testing. This study adds to the literature by determining if collaborative testing enhances students' understanding and lowers test anxiety in a high school setting.

**Literature Review**

*Assessment.*  Assessment is an essential part of education as it enables educators to collect information about students' academic learning, reasoning skills, and attitudes (Rao et al.,

2002). Quizzes and tests are the primary mechanisms used to determine if students have met the course's goals. The evaluation also determines student grades and their advancement into future classes (Giuliodori et al., 2008). Leight et al. (2012) describe how high stakes summative assessments like midterms, finals, and Advanced Placement (A.P.) exams help instructors know whether they have developed the required level of understanding concepts taught. They argue that testing students are the best method to ensure students retain the course material. Although this may be true, students may not understand what they did right or wrong as they do not see the exams. Therefore, students will not lbe able to determine if their knowledge or the strategies used were effective, reducing the assessment's pedagogical value (Leight et al., 2012; Rao et al., 2002). Additionally, individual testing does not consider disadvantages such as cultural differences, different learning styles, and additional challenges faced by English second language learners (Zapatero et al., 2012). Furthermore, traditional individual assessments do not consider social constructivist theories where students learn best in collaborative classrooms (Guiliodori et al., 2008; Zapatero et al., 2012).

*Collaborative learning.*  Social-cultural theories of learning and teaching, developed by Bruner and Vygotsky in the late nineteen sixties and seventies, construe that students are stimulated to learn and grow through social interactions (Seifert & Sutton, 2009). Constructivism, another learning theory, explains students actively construct their knowledge out of shared experiences, which augment their metacognition (Seifert & Sutton, 2009). Collaborative learning is a pedagogical method that utilizes social-cultural and constructivism to enhance student learning. This approach to education is powerful as students are actively engaged in their learning as they converse with their peers in small groups, exchanging and defending their ideas (Ghaith, 2018; Wanzek et al., 2014). Additionally, through shared learning, students expand their cognitive skills and ideas and develop new attitudes (Meseke et al., 2010).

Furthermore, research has shown many benefits to collaborative learning such as; additional academic and social support, increased student self-esteem, positively affected student achievement, increased student motivation, improved intergroup relations, improved critical and creative thinking, and improved problem-solving skills (Baloche & Brody, 2017; Leight et al., 2012; Meseke et al., 2010; Ngotngamwong, 2014; Rao et al., 2002). For collaborative learning to be effective in the classroom, the educator should be responsible for building teams. The teacher places the students in groups to ensure that they can improve their skills and develop their knowledge (Wanzek et al., 2014). Students retain information better and enhance their understanding of the concepts when they are not homogenous (Wenzel, 2000).

*Collaborative testing.*  Collaborative testing places the students in pairs or small groups of three or four for the test. Once in the group, the students can discuss the questions and then submit an individual test paper or a group test paper (Leight et al., 2012; Meseke et al., 2010; Weimer, 2018). There have been numerous research studies done, mainly with college students, into this assessment method's benefits. Giuliodori et al. (2008)

demonstrated an increase in collaborative test scores for high and low-performing students with the collaborative test groups. Albeit lower-performing students benefitted more.
In contrast, Gilley & Clarkston (2014) discovered that all students learned from the collaborative assessments irrespective of their performances. Seigel et al. (2015) undertook a study whereby the group testing occurred first, and then the students took the remainder of the test individually. They discovered that when the exam content is conceptually more straightforward, intermediate and low performing students benefit the most from group testing. Nonetheless, when the concepts are more complicated, all students benefit from collaborative testing. Thus this method enables students to utilize higher order thinking skills, which is a critical aspect of the N.G.S.S. (Gilley & Clarkston, 2014; Meseke et al., 2010).
 Students' level of retention of the material in collaborative testing has had mixed results. Gilley & Clarkston (2014) found greater retention and understanding of the students' concepts when they undertook group tests instead of individual tests. The students who attempted the group tests retained the information and correctly responded when taking a written pop quiz three days later. In contrast, students who did not participate in the group exam did not increase their retention or understanding of the material (Gilley & Clarkston, 2014).  Zipp (2017) undertook a collaborative assessment study where the students completed each exam individually and then were placed in groups to retake the exam. The results indicated that if the students answered correctly in the group test, they responded correctly on the final examination, two months later (Zipp, 2007). Other studies have also indicated that students who take group tests can longer retain the information (Cortright et al., 2003; Rao et al., 2002; Vogler & Robinson, 2016).

In contrast, Leight et al. (2012) and Sandahl (2010) discovered no greater retention occurred between the control group of individual test-takers than the experimental group of collaborative test takers on the final exam. An explanation for this difference may be that students who have had more exposure to the content have a sufficient prior understanding to enable new knowledge to be assimilated and retained (Leight et al., (2012). Despite not showing a difference in retention in the final, Leight et al. (2012) results indicate that cooperative testing increased students' level of engagement and confidence and enhanced their understanding of the concept.

With a better conceptual understanding and retention, collaborative classrooms have many benefits for students, including enhanced participation, increased social skills, and higher critical thinking and problem-solving skills (Baloche & Brody, 2017; Ngotngamwong, 2014). Assessment that uses collaboration between students should then have positive effects on students. Despite not showing a difference in retention in the final, Leight et al. (2012) indicate that cooperative testing increased students' level of engagement and confidence and enhanced their understanding of the concept. Likewise, Ngotngamwong's (2014) study determined most students (over 80%) felt that pair testing was enjoyable, created more outstanding teamwork and cooperation between the students, and ensured they studied harder. Hanshaw (2012) and Rao et al. (2002) and others describe many benefits from this testing method: positive influence on learning, better interpersonal, communication, conflict resolution, and critical thinking skills. Additionally, their results indicated increased metacognitive skills, increased persistence to problem solve, enhanced memory and

retention, and effective listening skills, all vital skills for students (Gilley & Clarkston; Hanshaw 2012; Kapitanoff, 2009; Rao et al. 2002).

Coupled with the above benefits, collaborative testing has been shown to diminish test anxiety in students, as described in more detail below (Breedlove et al., 2004; Cortright et al., 2003; Krispenz & Dickhäuser, 2018).  Some studies indicate there are negative consequences to collaborative testing. A typical adverse claim is that students are not preparing for the assessment (Giuliodori et al., 2008). One method used to alleviate this is to ensure the students do not know they are undergoing a collaborative test until they arrive at the classroom (Rao et al., 2002). Moreover, some students have said that they could not reach a consensus on the answers or had mismatched partners (Ngotngamwong, 2014; Zipp, 2017). To overcome this effect, students could hand in an individual copy of the assessment and consequently do not need to reach a consensus (Ngotngamwong, 2014). *Test Anxiety.*  Test anxiety relates to students' emotions when studying for and taking an exam (Krispenz & Dickhäuser, 2018). Test anxiety can have many consequences for students, including a cognitive component whereby the student cannot retrieve the information, the students may have difficulties in organizing and retaining information, or the students feel overcome by a fear of failure (Breedlove et al., 2004; Krispenz & Dickhäuser, 2018). The physiological reactions of test anxiety may comprise trembling, palpitations, sweating, dizziness, and nausea, impacting student well-being (Breedlove et al., 2004; Krispenz & Dickhäuser, 2018). Test anxiety can also affect the students' mental health as it can lower self-esteem, increase feelings of helplessness and insecurity, erode confidence, and diminish motivation (Breedlove et al., 2004; Krispenz & Dickhäuser, 2018). Equally important is the knowledge that test anxiety impacts student academic achievement, affecting each student's educational and employment prospects (Krispenz & Dickhäuser, 2018).

Multiple studies using collaborative assessments have demonstrated that this assessment method reduces test anxiety in students (Breedlove et al., 2004; Leight et al., 2012; Meseke et al., 2010; Pandey, C. & Kapitanoff, S., 2011). Reduction in test anxiety may be brought about by students sharing prior learned information, discussing questions and answers, and having intellectual support from their peers (Breedlove et al., 2004; Rao et al., 2002).

**Methodology**

*Statement of Purpose.*  This action study was undertaken in an Advanced Placement Environmental science classroom after the researcher noticed students benefited when undertaking collaborative work. The students described this method as enabling them to get different perspectives and develop their understanding of the concepts. Moreover, it was observed by the researcher that students felt very stressed when undertaking summative tests, and hence some students were not able to perform to their ability. Therefore, the researcher wanted to determine if collaborative testing could be a method to increase understanding and reduce test-taking anxiety in a high school.

The benefits of assessment in groups have been studied numerously over the years, generally in college science classes, as cited in the studies described above. However, for high schools, there have been very few studies undertaken. Therefore, this current study was conducted to determine if the impact of collaborative assessments in a high school was similar to the collegiate findings, including academic performance and reduction of test anxiety. This study, therefore, was guided by the following research question:

- Does collaborative testing in a high school A.P. Environmental Science course positively impact student test outcomes?
- Does collaborative testing in a high school A.P. Environmental Science course lower student test anxiety?
- How do students view collaborative testing in terms of preparing them for the A.P. exam?

*Course Structure and Content.*  This study occurred over the fall, winter, and spring semesters of the 2018 -2019 school year. The participants included forty-one high school students in two sections of an Advanced Placement Environmental Science course taught by different teachers.  The student population for this study consisted of thirty-eight seniors and three juniors (N=41), of which sixteen (39%) were males, and twenty-five (61%) were females. Four students missed a collaborative test; therefore, all of their testing data were excluded from analysis (N=37), but their survey results were still analyzed.

The A.P. Environmental Science is a full-year course, meeting for four sixty-minute periods per week. These periods consisted of a mix of lectures, collaborative laboratory assignments, and inquiry-based work. During November, the students participated in a group quiz on Renewable Energy, allowing them to be familiar with undertaking a collaborative examination. The groups were teacher assigned based on previous test scores, gender, and grade level. The makeup of these groups is outlined in Appendix A. During this research period, the topics covered were Renewable Energy, Climate and Biomes, Indoor and Outdoor Pollution, Climate Change, and Agriculture and Food. Summative assessments, which were identical for both classes, for each topic, were made up of 21-25 multiple choice questions, each with five answers and four to six short answer questions (20 minutes). The questions were of varying complexity and sourced from previous A.P. Environmental Science examinations run by the College Board, U.S.A.

*Research Design.*  This study was undertaken to determine if collaborative testing could improve student achievement in a high school APES class. To accomplish this, the students end of topic tests were used to generate data to compare the individual and collaborative test results. Students attempted three individual and three collaborative tests, enabling a comparison between the student's individual and collaborative testing grades (Giuliodori et al., 2008). The classroom teacher determined the heterogeneous testing groups' assignments based on students' prior academic achievement, gender, motivation, and ability to stay on task (Wanzek et al., 2014). To ensure randomness, the students were unaware of their group assignments until entering the testing room (Meseke et al., 2010). Due to the small class size, the groups consisted of three or four participants (See Table 1). Additionally, because of gender inequality in the study group, there were more females in

some groups than males; however, every group included at least one female and one male. Table 1 describes the number of participants in each group for collaborative tests.

*Table 1:  Test topic and group makeup*

| Test Topic | Number of groups | Group makeup | Total Students | Students in analysis |
|---|---|---|---|---|
| Renewable energy | 13 | 11 groups of 3<br><br>Two groups of 4 | 41 | 37 |
| Climate and Biomes | Individual test | | 41 | 37 |
| Indoor pollutants | Individual test | | 41 | 37 |
| Climate change | 13 | 8 groups of 3<br><br>4 groups 4 | 40 | 37 |
| Outdoor Pollutants | Individual test | | 41 | 37 |
| Agriculture and Food | 13 | 9 groups of 3<br><br>3 group 4 | 39 | 37 |

*Note. The groups were determined by gender, previous test scores, and class rank.*

The collaborative group test was undertaken under the same conditions as the individual test to determine any improvements between individual and collaborative testing (Wanzek et al., 2014). The students were allotted 50 to 55 minutes to answer the questions. Students did not know if a test would be collaborative or individual until they were in the classroom, ensuring they studied for the test. Each group submitted their own exam paper with all of the group names on the document (Leight et al., 2012; Nanzek et al., 2014; Wanzek et al., 2014). Moreover, the students submitted their test papers in the collaborative testing phase, thus enabling them to change their responses if they could not reach a consensus on an answer (Ngotngamwong, 2014). To ensure grading consistently between the teachers the open-ended questions were graded using the college boards APES rubric for each test.
To answer the second quantitative question about whether collaborative testing affects students' test anxiety, they undertook an anonymous online survey distributed through Qualtrics two days after the final summative assessment. The survey questions included collaborative testing and test-taking anxiety (See Appendix B for survey questions) adapted for high school students from questions in Cortright et al. (2003), Hanshaw (2012), and

Ngotngamwong's (2014). Lastly, the researcher undertook an anonymous survey to determine if the students felt that collaborative testing aided their retention when undertaking their final exam (Appendix C for survey questions).

*Data Analysis.* To answer research question 1, descriptive statistics including averages from percentage scores and standard deviations were calculated for all six exams. The data collected only included students present at testing for both individual and collaborative tests (N=37). The data gathered was analyzed by a one-way repeated measures ANOVA test and paired-sample t-tests through IBM SPSS to ascertain the student differences between group and individual testing. The assessments' results were examined in two ways; the first was a repeated measure ANOVA that looked at the differences between the three independent and the three collaborative tests. The second test was a paired t-test to compare the averages of the individual and group tests. Additionally, a paired t-test was undertaken to investigate whether there were overall differences between the two tests' averages (Shier, 2004).

After the testing period ended, students were asked to complete an online survey (modified from Hanshaw, 2012 & Meseke et al., 2010) to determine their perceptions of collaborative testing on understanding and whether it affects their test-taking anxiety. The survey asked 19 questions and was based on the 0-5 Likert scale (Leight et al., 2012). All students(N=41) undertook the survey; a subset of the responses is represented in Table 4. Lastly, after the A.P. exam in May, students were informally questioned to determine if collaborative testing affected their understanding and retention of the concepts. Their responses are listed in Table 5.

**Results**

*Student Achievement.* For the three independent tests, the means (See Figure 1) are as follows with the standard deviation in parentheses, 78.22 (11.15), 77.41 (13.14), 78.62 (11.20). The Mauchly's sphericity tests had been met, as detailed in Table 2 (df = 2, Sig. = 0.856), $F (2, 72) = 0.156$, $p = 0.64$, indicating there is not a significant difference between the three individual test scores. Regarding the collaborative tests, there was a significant difference in the means (See Figure 1), 85.08 (11.63), 89.73 (5.47), and 94.19 (5.76), further validated by the results of the ANOVA test, whereby Mauchley's test of sphericity indicate there is a significant difference between the collaborative tests ($p < 0.000$). Due to the sphericity not being met, a Greenhouse-Geisser correction was applied (Table 2).

*Table 2: ANOVA Data for Independent and Collaborative Tests*

| Variable | N | Mauchley's Sphericity | df | F | p | Greenhouse-Geisser |
|---|---|---|---|---|---|---|
| Independent | 37 | 0.640 | 2 | 0.156 | 0.64 | |

| Collaborative | 37 | 0.000 | 1.448 | 11.379 | 0.001 | 0.00 |
|---|---|---|---|---|---|---|

*Note. The Greenhouse -Geisser test was added as the collaborative test did not meet the standard for sphericity.*

To determine if there was a difference between the averages of the independent and group test scores, a paired t-test was undertaken. The t-test was between the group and individual test 1, group and individual test 2, and group and individual test 3. The results are displayed below in Table 3. There is a significant difference in test scores between collaborative and individual testing. This is also depicted in the difference in means between the tests and illustrated in Figure 1. The box plot shows that students who undertook collaborative testing generally had increased scores compared to their individual test scores.

*Table 3. Paired t-test between the averages of the three individual and group tests*

| Variable | Mean | Std. Deviation | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|
| Pair 1- ind. Test 1- gp 1 | -6.86 | 16.13 | -2.588 | 36 | .014 |
| Pair 2 – ind Test 2 - gp 2 | -12.32 | 13.38 | -5.598 | 36 | .000 |
| Pair 3 – ind Test 3 - gp 3 | -15.57 | 12.21 | -7.756 | 36 | .000 |
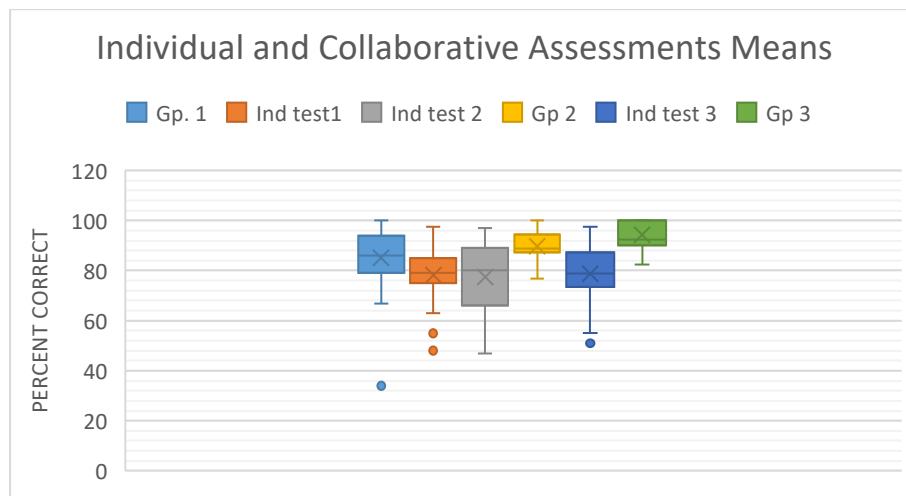


*Figure 1. Individual (Ind test) and group (Gp) test results from the A.P. Environmental Science Course were displayed when they took each test.*

Figure 2 below displays the averages of the three individual and collaborative tests for each student. Along with the t-test data, the graph indicates an overall increase in student test scores compared between individual and group tests, which is very apparent in individuals 3, 9, 25, 27, 30, and 36.  Although individuals 7, 8, and 10 had slightly better individual test scores than group test scores.
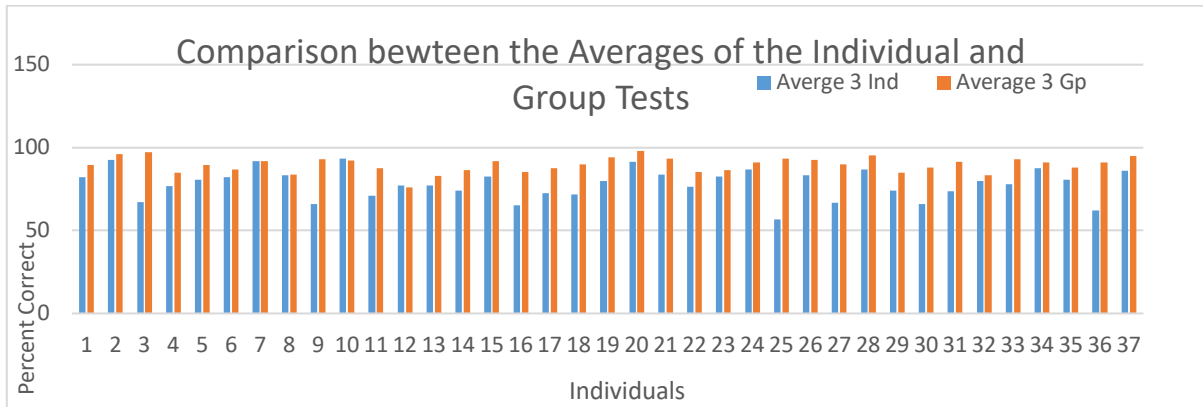
*Figure 2. Comparison of each student's average individual and collaborative percent test score.*

*Student Test Anxiety and Preparation for A.P. Exam.* To answer the second research, question all students (N=41) in the course were anonymously surveyed. The survey questions and responses are described below and in Table 4. 56% of students felt that individual testing increased their anxiety, whereas 26% said individual testing did not impact their anxiety levels. The following students' responses to the open-ended question represent most students' reasons for increased anxiety on the individual tests:

- It was very stressful because even if I knew the right answer, I second guessed it.
- Doing an individual test definitely increased my anxiety because I had no one backing up my answer or contradicting me on why the answer I picked was wrong.
- I felt that taking a test individually stresses me out more because I could be the only one getting the bad grade.
- I feel more anxious during individual tests.
- It is more stressful to take the individual section since you cannot discuss and get the opinions of your classmates.

On the other hand, the majority of students (87.88%) felt that undertaking collaborative testing lowered their anxiety levels, reasons for this from the open-ended questions included: regarding confidence in the group's answers, ability to discuss the questions, and understanding if they were not sure some member of the group would have the answer. The majority (81.81%) of students felt that working in a group was less stressful than undertaking individual tests. The student's comments below provide some of their reasoning:

- My group members were cooperative, we were all respectful and kind to each other, and if we disagreed, we just put a separate answer, and I get to interact with people and understand the material a lot more.
- I didn't feel very anxious after the test because if I didn't know the answer to a question someone else in my group did, which helped me feel like we as a group would get a good grade. I also felt that because it was a group test I wasn't the only getting the grade if it turned out to be bad."
- I feel as if it was less stressful since it was a group test and that I was more confident in my answers and more relaxed about getting my grade back.

9.09% of students said that it increased their anxiety levels; their reasons were due to always being stressed before a test. Besides, as one student states,

Sometimes they disagreed with me, and I didn't want to write something different, My group was divided over many questions, and neither side was willing to accept that they may be wrong, and I find myself being able to work to a more efficient degree when coupled with my peers; however, I am often look towards as the "intelligent" one of the group, and I feel slightly stressed due to the position of power I am involuntarily elected to.

*Table 4. Survey of student's anxiety levels related to individual and collaborative tests (N=41).*

| Questions | Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree |
|---|---|---|---|---|---|
| I felt anxious after individual test | 14.6% | 12.2% | 17.1% | 31.7% | 24.4% |
| I felt anxious after the group test | 61% | 24.4% | 7.3% | 4.9% | 2.4% |
| Group tests increased my anxiety | 73.2% | 14.6% | 2.4% | 9.8% | 0 |
| I found working in a group stressful | 46.3% | 36.6% | 4.9% | 4.9% | 7.3% |
| Working in a group helped my understanding of the content | 4.9% | 2.5% | 7.3% | 39% | 46.3% |

Lastly, the students were asked if collaborative testing increased their understanding of the content. Most students (87.87%) felt that their knowledge of the content increased, while 9.09% indicated it neither helped nor hindered their understanding, and one student thought it negatively impacted their ability. Representative student comments include:

- Talking about the problems helped me understand them better.
- Combining the knowledge of everyone increased my knowledge on different topics.
- By working with a group, we were able to discuss our reasoning and come to a correct answer.
- By bouncing our ideas off of one another, I felt all of the previous information I ascertained re-enter my mind and reposition itself into a more organized whole.

After the A.P. exam, 63.3% of students felt that this method of assessment aided their retention of the material, while 33.3% felt that it only helped their retention a little. Further, 66.7% of students stated a difference in retaining the information from the different testing methods (See Table 5).

*Table 5: Group testing and retention of material on the A.P. exam. (N=31)*

| Did you feel that you retained the information that was tested in the group tests? | Definitely | A little | No |
|---|---|---|---|

| | | |
|---|---|---|
| | 19 (63.3%) | 10 (33.3%) | 2 (6.7%) |
| Did you feel there was a difference in retention between the material from the group or individual tests? | Yes | No | Maybe |
| | 66.7% | 6.6% | 26.7% |

*Note*. This survey was given post-A.P. exam, only 31 students surveyed as the remaining were not in class at the time.

**Limitations**

The amount of time for the parental approvals to be returned impacted the time available to undertake the tests. In turn, this caused the researcher to have to speed up the testing process. Hence, students could determine when the collaborative or individual tests would be undertaken before walking into the room. It could not be as random as initially planned. Some students were also removed from the testing data but not survey data because the survey was undertaken anonymously, and the researcher did not know the individual student responses.

Another limitation may include that the study was conducted in two science classrooms with different teachers. Therefore, the other class students may have known when the group tests were going to be undertaken and the content. Knowing the test's content could influence the test scores as the students understand what subject matter to focus on, leading to inflated test scores. Additionally, knowing when the group tests were undertaken could impact the student's motivation to study. This, in effect, did occur as some students, after the last test, did indicate to the researcher that knowing the content and that there was to be a group test caused them to study less and rely more on other people in the groups. Another limitation would be biased when responding to the questions; this could reduce the results' reliability. Due to the students in the survey and last informal questioning responding with answers they think the researcher wanted to have. Lastly, this study did not measure the effects of collaborative testing on retention. Retention was informally discussed with the students.

**Discussion**

The purpose of this study was to determine whether there were differences between individual and collaborative test scores and levels of test anxiety in students undertaking an A.P. environmental science course. This study's findings support the thesis that most students perform better when undertaking collaborative rather than individual tests (Table 3), as shown in previous research (Giulidori et al., 2008; Haberyan & Barnett, 2010; Rao et al., 2002). The individual repeated measures test means were not substantially different from each other (p = 0.64), indicating there were no significant differences between the mean student scores of the three individual tests. Regarding the group tests, all three means were higher than the individual test means, although the third group test had a significantly

higher mean (94.19) than the previous two group tests (85.08 and 89.73, p <0.001). There are various reasons as to why this increase may have occurred. The students in this study had not undertaken collaborative testing previously; therefore, it may have taken them some time to become accustomed to working in a group setting. Moreover, the dynamics of the different groups may have impacted the student's ability to stay on task in the first collaborative test, as they may have been distracted by working with friends or, conversely, with individuals they did not know very well. Additionally, as the students became accustomed to collaborative tests, they may have felt that if they undertook some effort, then as a group, they all could do well on the assessment.

Furthermore, when comparing the individual test to the group test, the results demonstrate a considerable difference between the two testing methods. The group tests had a significant average increase (See Table 2 and Figure 3). These and prior study results lead us to conclude that collaborative testing can facilitate student learning (Leight et al., 2012; Ngotngamwong, 2014; Rao et al., 2002; Vogler & Robinson, 2016). Collaborative testing can facilitate learning as students analyze and discuss the questions and explain their reasoning to each other. This knowledge sharing can lead to greater understanding (Rao et al., 2002). Many of the students in our survey indicate this was the case; they felt that by discussing the questions, they understood the concepts in greater depth, similar to Siegel et al. (2015) findings.

Moreover, improved testing scores could be due to students stimulating their thinking through activating prior understanding. When discussing each question, students remember what they have learned in the past related to the problem, thereby building their knowledge (Leight et al., 2012; Siegel et al., 2015). This was evident in the student's responses to the survey question regarding whether group testing enhanced their understanding of the concepts. Many students stated that discussing the group's problems gave them a deeper understanding of the ideas (See survey comments). While working in their groups, most students were engaged in discussions and felt they contributed equally. However, as noted in the student comment above, a student felt the group relied on one individual. The reliance on specific individuals to "carry the team" has been noted in other research (Rao et al., 2002; Seigel et al., 2015). In this study, as with Rao et al., (2002) study, the researcher tried to alleviate this concern by randomly selecting the individuals in each group based on the criteria mentioned above. Collaborative testing can enhance student motivation to study. In this paper, students were not told what type of test they would undertake until they were in the classroom; therefore, the researcher surmises the students studied how they typically would for a test. However, this study did not test whether students would study more or less if they were made aware of the test structure beforehand.

The survey results agree with many other researchers that collaborative testing can lead to lower anxiety levels in students (e.g., Cantwell et al., 2016; Pandey & Kapitanoff, 2011; Seigal et al., 2015). In this study, many students reported that undertaking individual tests increased their anxiety, however, most felt this was negated by working with their peers (see table 4 and student comments).

Student beliefs in how they perform on tests can negatively impact their ability to access their working memory and lower their grades (Krispenz & Dickhäuser, 2018). The students reported that working in groups enabled them to feel that if they had forgotten a fact, the other team members would know the answer; however, when working individually, they were not able to discuss the solutions, and therefore their anxiety levels would increase pre and during the test. Concerning the collaborative tests, 87% of students felt that working in collaborative groups significantly lowered their anxiety. The lessening of test anxiety documented by most students may have enabled students to focus more on the questions at hand rather than how they will perform on the test, thus increasing their participation and consequent grades (Hanshaw, 2012).

There are some concerns related to collaborative testing. Siegel et al. (2015) and Haberyan & Barnett (2012) found that students felt some groups were more unequal in ability and group dynamics. In this study, the groups changed each test, and the students did not know before the test which group they were going to be placed in. Nevertheless, during in-class observations, some groups did not work effectively, as is evident by the negative comments above. This may be due to the dynamics of high school senior social interactions and differences in personalities (Haberyan & Barnett, 2012).
Furthermore, "social loafing" may have occurred, where one or more group members do not participate in the discussions, which may lessen this assessment method (Rao et al., 2002). Social loafing was seen in this study by a few students, as mentioned in the comments above. Nonetheless, most students (89.7%) felt they each member contributed equally to the group. Another concern relates to students who understand the material in-depth compared to other students, leading to an inflation in grades (Cantwell et al., 2016; Siegel et al., 2015). Figure 2 illustrates how some students' grades increased more significantly than others compared to their tests. In this study, there were five students (12%) with notable grade increases. Grade inflation may impact the higher-performing students as their grades do not increase as much as the lower performing students. Additionally, if the students use these grades to then move into more challenging classes, they may lack the ability to perform well (Cantwell et al., 2016). Garde inflation can be mitigated by ensuring no more than 15% total worth is assigned for the collaborative test section. Moreover, Cantwell et al., (2014) proposed that if a student fails the individual portion, they do not have their group scores added or did not participate.

A more significant issue related to group testing that was not quantitatively addressed by this researcher was student retention and whether this impacts more comprehensive exams. A significant portion (96.6%) of students indicate that collaborative testing helped them retain the information. However, the literature is divided as to whether this method of assessment aids student retention. Some studies, such as those undertaken by Bloom (2009) and Cortright et al. (2003), note some improvement in students' retention. On the other hand, Leight et al. (2012) and Sandahl's (2010) studies indicate little retention by their students. Further research is required to ascertain whether collaborative testing enhances retention in students.

**Conclusion**

Collaborative testing provides an alternative method to assess students and has had positive results in this and other studies (Breedlove et al., 2004; Leight et al., 2012; Meseke et al., 2010; Seigel et al. 2015). This method positively impacted student test outcomes for most students in the study and provided students with opportunities to think and cultivate different viewpoints critically. Moreover, results from the survey on group assessments indicate that students' test-taking anxiety was reduced.

Currently, in secondary education, there is a reform movement whereby collaborative learning practices are being utilized more often to promote deeper engagement and understanding of the concepts (Nariman and Chrispeels, 2016). However, minimal collaborative testing is attempted in high schools. As an educator, the researcher believes this assessment method should be utilized more often in secondary education. The academic and social benefits of collaborative assessments for high school students are many. This method allowed the students to discuss questions. They had to defend their opinions and listen to other perspectives. In doing so, the students reasoned through the questions, leading them to understand the concept. Furthermore, collaboration in heterogenous groups enabled most students to lessen anxiety related to test-taking, which can negatively impact student achievement. The issues concerning group testing are minor when compared to the overall benefits. In the future, the researcher will use this method as an assessment tool with a few modifications. The groups should be chosen from a larger cohort to ensure more heterogeneous groups, in different science classes.  Additionally, the researcher believes that pacing the tests throughout the year would provide more accurate test results.

**About the Authors**

**Helen Bremert** is currently undertaking a Ph.D. at Sydney University in the School of Education and Social Work, with the topic on assessment.  She recently completed a Master's in Chemistry Education at Illinois State University.  Prior and during her recent graduate work, she was a science teacher in New Jersey, and Sydney, Australia. Email: hbre8322@uni.sydney.edu.au

**S. B. Boesdorfer, Ph.D.** is an assistant professor of chemistry education in the Department of Chemistry at Illinois State University in Normal, Illinois. Email: sbboesd@ilstu.edu

## References

Baloche, L. & Brody, C. M. (2017). Cooperative learning: Exploring challenges, crafting innovations. *Journal of Education for Teaching*, 43(3), 274-283. doi:10.1080/02607476.2017.1319513

Breedlove, W., Burkett, T., & Winfield, I. (2004). Collaborative testing and test anxiety. *The Journal of Scholarship of Teaching and Learning, 4*(2), 33–42.

Balkam, B., Nellessen, J., & Ronney, H. (2013). Using Collaborative testing to reduce test anxiety in elementary and middle school students, Unpublished manuscript, Retrieved from https://files.eric.ed.gov/fulltext/ED541337.pdf

Cantwell, R., Sousou, J., Jadotte, Y., Pierce, J., & Akioyamen, L. (2017). Collaborative testing for improving student learning outcomes and test-taking performance in her education: A systematic review. *Campbell Systematic Reviews, 13*(1), 1–18. https://doi.org/10.1002/CL2.186

Efron, S. A., & Ravid, R. (2013). Action research in education: A practical guide. The Guilford Press.

Ghaith, G. M. (2018). Teacher perceptions of the challenges of implementing concrete and conceptual cooperative learning. *Issues in Educational Research, 28*(2), 385–404.

Gilley, B., & Clarkston, B. (2014). Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. *Journal of College Science Teaching, 43*(3), 83-91. https://doi.org/10.2505/4/jcst14_043_03_83

Giuliodori, M. J., Lujan, H. L., & DiCarlo, S. E. (2008). Collaborative group testing benefits high- and low-performing students. *Advances in Physiology Education, 32*(4), 274-278.  https://doi.org/10.1152/advan.00101.2007 PMID:19047503

Haberyan, A., & Barnett, J. (2010). Collaborative testing and achievement: Are two heads really better than one? *Journal of Instructional Psychology, 37*(1), 32–41.

Hanshaw, L. G. (2012). Qualitative aspects of group-only testing. *College Student Journal, 46*(2), 419-426.

Kapitanoff, S. H. (2009). Collaborative testing: Cognitive and interpersonal processes related to enhanced test performance. *Active Learning in Higher Education,10*(1), 56–70. https://doi.org/10.1177/1469787408100195

Krispenz, A., & Dickhäuser, O. (2018). Effects of an inquiry-based short intervention on state test anxiety in comparison to alternative coping strategies. *Frontiers in Psychology, 9*(201), 201. https://doi.org/10.3389/fpsyg.2018.00201 PMID:29515507

Kurt, A. S., Balci, S., & Kose, D. (2014). Test anxiety levels and related factors: Students preparing for university exams. *JPMA. The Journal of the Pakistan Medical Association, 64*(11), 1235–1239. PMID:25831637

Leight, H., Saunders, C., Calkins, R., & Withers, M. (2012). Collaborative testing improves performance but not content retention in a large-enrollment introductory biology class. *CBE Life Sciences Education, 11*(4), 392–401. https://doi.org/10.1187/cbe.12-04-0048 PMID:23222835

Meseke, C. A., Nafziger, R., & Meseke, J. K. (2010). Student attitudes, satisfaction, and learning in a collaborative testing environment. *The Journal of Chiropractic Education, 24*(1), 19–29. https://doi.org/10.7899/1042-5055-24.1.19 PMID:20480011

Nariman, N., & Chrispeels, J. (2015). P.B.L. in the era of reform standards: Challenges and benefits perceived by teachers in one elementary school. *The Interdisciplinary Journal of Problem-Based Learning, 10*(1). Advance online publication. https://doi.org/10.7771/1541-5015.1521

NGSS Lead States. (2013). Next Generation Science Standards: For States, By State*s*. The National Academies Press.

Ngotngamwong, R. (2014). Pair tests in a high school classroom. *SAGE Open, 4*(4),1-8. https://doi.org/10.1177/2158244014564359

Nist, P., & Diehl, M. (1990). PHCC test anxiety questionnaire. Retrieved from http://phcc.edu/ods/questionnaire.html

Pandey, C., & Kapitanoff, S. (2011). The influence of anxiety and quality of interaction on collaborative test performance. *Active Learning in Higher Education, 12*(3), 163–174. https://doi.org/10.1177/1469787411415077

Rao, S. P., Collins, H. L., & DiCarlo, S. E. (2002). Collaborative testing enhances student learning. *Advances in Physiology Education, 26*(1-4), 37–41. https://doi.org/10.1152/advan.00032.2001 PMID:11850326

Reiser, E. (2017). Blending individual and group assessment: A model for measuring student performance. *The Journal of Scholarship of Teaching and Learning, 17*(4), 83–94. https://doi.org/10.14434/josotl.v17i4.21938

Ross, K. N. (1978). Sample design for educational survey research. Pergamon Press. https://doi.org/10.1016/0145-9228(78)90001-8.

Sandahl, S. S. (2010). Collaborative testing as a learning strategy in nursing education. *Nursing Education Perspectives, 31*, 142–147. PMID:20635617

Seifert, K., & Sutton, R. (2009). Educational Psychology. Retrieved from https://courses.lumenlearning.com/educationalpsychology/chapter/major-theories-and-models-of-learning/

Shier, R. (2004), Statistics: 1.1 Paired t-tests. Retrieved from http://www.statstutor.ac.uk/resources/uploaded/paired-t-test.pdf

Siegel, M., Roberts, T., Freyermuth, S., Witzig, S., & Izci, K. (2015). Research and teaching: aligning assessment to instruction: Collaborative group testing in large-enrollment science

classes. *Journal of College Science Teaching, 44*(6), 74–
81. https://doi.org/10.2505/4/jcst15_044_06_74

Suinn, R. M. (1969). The STABS, a measure of test anxiety for behavior therapy: Normative
data. *Behaviour Research and Therapy, 7*, 335–339. https://doi.org/10.1016/0005-
7967(69)90018-7PMID:5344977

Test Anxiety Self-Assessment (N.D) - nhti.edu. Retrieved
from http://www.nhti.edu/sites/default/files/content/documents/lctestanxiety.pdf

Vogler, J. S., & Robinson, D. H. (2016). Team-based testing improves individual learning. *Journal
of Experimental Education, 84*(4), 787–
803. https://doi.org/10.1080/00220973.2015.1134420

Wanzek, J., Vaughn, S., Kent, S., Swanson, E., Roberts, G., Haynes, M., & Solis, M. (2014). The
effects of team-based learning on social studies knowledge acquisition in high
school. *Journal of Research on Educational Effectiveness, 7*(2), 183–
204. https://doi.org/10.1080/19345747.2013.836765

Wenzel, T. J. (2000). Cooperative student activities as learning devices. *Analytical Chemistry, 72*,
293A–296A. https://doi.org/10.1021/ac0027643 PMID:10763228

Wieman, C. E., Rieger, G. W., & Heiner, C. E. (2014). Physics exams that promote collaborative
learning. *The Physics Teacher, 52*(1), 51–53. https://doi.org/10.1119/1.4849159

Zabrucky, K. M., & Bays, R. B. (2015). Improving student's retention of classroom material
through the testing effect. *College Teaching, 63*(2),
91. https://doi.org/10.1080/87567555.2015.1005043

Zapatero, E. G., Maheshwari, S. K., & Chen, J. (2012). Effectiveness of active learning
environment: Should testing methods be modified? *Academy of Educational Leadership
Journal,16*(4), 101–114.

Zipp, J. F. (2007). Learning by Exams: The Impact of two-stage cooperative tests. *Teaching
Sociology, 35*(1), 62–76.  https://doi.org/10.1177/0092055X0703500105

## Appendix A:  Seating Groups for Group Tests

**First group test**
HB (5-13-17) (1-24-25) (18-21-31-39) (5-34-40) (32-38-28) (26-7-29)
CB (19-37-8) (35-4-14) (22-2-3) (33-11-27) (37-36-9) (20-12-23)

**Second group Test**
HB  - (28-5-1) (31-24-30) (25-6-29) (15 – 34 – 39 – 13) (7 – 38 – 32) (18-26-40)(21-17-36)
CB - (22-3-12-27) (35-4-14-8) (9-19-2) (37-20 – 16) (10-11-33-23)

**Third group test**
HB (28-31-30) (1-40-5) (24-38-39) (25-13-32) (7-17-15) (34-26-6)
CB (3-22-23) (4-8-12) (20-37-35-11) (33-2-14-19) (36-9-27-10)

(16, 18, 29, 21) have been removed as absent for a group test)

**Appendix B:  Post Test Survey Questions**

1. How many hours did you study for this test?
2. I felt confident I was properly prepared for the test
3. I felt positive during the test
4. The individual test section increased my anxiety – then explain choice
5. The group test section increased my anxiety – then explain choice
6. I felt more relaxed working in a group
7. I felt relief when undertaking the individual portion of the test
8. Working in a group helped my understanding of the content – then explain my choice
9. I found working in a group stressful – then explain my choice

**Appendix C:  Post A.P Exam Survey Questions**

1. Did you feel there was a difference in retention between the material from the group or individual tests?
2. Please explain your answer
3. Did you feel that you retained the information that was tested in the group tests?
4. Please describe the reason for your answer.

# INVESTIGATING CO-TEACHING FOR IMPACT ON ACADEMIC ENGAGEMENT: BEST PRACTICES FOR ENGLISH SUBJECT LEARNERS IN A BILINGUAL ELEMENTARY SCHOOL

**Kyle M. Daley**
**University of West Florida**

**Abstract** While the benefits of co-teaching have been espoused by many, the literature shows a lack of data supporting its effectiveness in achieving student learning outcomes and improving engagement. Using a convergent mix-methods design this study investigated the effectiveness of co-teaching on student engagement by observing a 3rd grade English class in a bilingual-school setting. Through several observations the study measured the self-reported engagement of students in co-taught and non-co-taught lessons, as well as the class teacher's perspective on student engagement. The investigation found a lack of concrete support for co-teaching's effectiveness for improving student engagement.

## Introduction

The concept of co-teaching as an instructional model has developed out of the ideas surrounding inclusive classrooms and special education since the 1960's (Aliakbari & Nejad, 2013; Cook & Friend, 1995; Pappamihiel, 2012; Yopp et al., 2014). Today, this instructional strategy is an increasingly popular method within both special education and general education classrooms. The increased use of co-teaching can be linked to statutes found in the Individual with Disabilities Education Act supporting the concept of Least Restrictive Environment (LRE) and provisions regarding appropriate interventions prior to student referral for special education evaluation (Baca, 1990; Oh et al., 2017). In the general education classroom, it is seen as a means to best utilize teacher strengths, improve student motivation, vary instructional methods, and increase student learning (Magiera et al., 2006). Moreover, it is linked with other positive benefits for both teachers and students, such as lower student to teacher ratios, more instructional support, and more peer-to-peer learning.

However, there are shortcomings in the literature regarding the lack of concrete data supporting co-teaching's effectiveness on student learning outcomes and engagement (Alikbari & Nejad, 2013; Yopp et al., 2014). There also appears to be a lack of general research concerning how co-teaching can be best used with English Language Learners (ELLs) and their needs as a group requiring support and differentiation (Aliakbari & Nejad, 2013; Maryland Department of Education, 2012; Pappamihiel, 2012; Magiera et al., 2006). The literature review will look at the origins of the co-teaching methods, their most common manifestations, and gaps within the subject. Based on the limited data supporting the effectiveness of co-teaching on engagement and with a multilingual student body, the following question is proposed:  How does co-teaching impact student engagement compared to non-co-teaching in a multilingual student environment?

**Literature Review**

*Co-Teaching Strategies.* Much of the literature agrees that each specific co-teaching method should adhere to four tenets (Cook & Friend, 1995; Pappamihiel, 2012). To be considered true co-teaching, the following cases must be met: 1) instruction is given by two professional educators, 2) both teachers must deliver instruction, 3) the classroom must be diverse and 4) the students must remain in one dedicated space. Since the method originated in the special education field, most of the literature references a pairing of a general education teacher and a special education teacher (Cook & Friend, 1995; Maryland Department of Education, 2012; Pappamihiel, 2012). The special education teacher or specialist originally acted in a supporting role to assist specific students. However, as co-teaching has become more commonplace in the mainstream classroom, where support is generalized and not necessarily targeted at specific students, teacher pairs may include any type of subject or specialist combination. In effect, the pairing of the professionals should reflect the diverse needs of the class and support inclusion for all students (Cook & Friend, 1995; Taşdemir & Yıldırım, 2017).

In terms of the instructional methods, the literature generally agrees that there are five co-teaching methods. None of the methods are deemed superior to any other, but in fact each would be more appropriate for certain types of lessons or tasks (Brendle et al., 2017; Chang, 2018; Cook & Friend, 1995; Taşdemir & Yıldırım, 2017). In each manifestation the two educators assume different roles of interaction with the class and one another. Cook and Friend (1995) outlined the strategies, starting with the One Teaching-One Assisting or Teacher Floater model. In this first of five arrangements, one teacher conducts the class while the other collects data or assists students. This first model is the most commonly used by teachers as it requires the least preparation, the least knowledge or experience with co-teaching methods, and it is often felt by teachers to be the most comfortable arrangement in terms of role identification (e.g. both teacher understand their roles and related expectations) (Magiera et al., 2006). Next, there is Station Teaching, where instructional content is divided into specific topics and organized in small group sections around the classroom.  With this model teachers operate sections independently, so there is little issue with balancing teachers' roles in the classroom. However, there is a greater level of preparation needed to organize the lesson as well as the classroom. Furthermore, in this

model there is normally at least one group of students working independently. The third model is referred to as Parallel Teaching. In this model the class is divided in half with each teacher delivering the same content material. This method allows for better use of individual teaching styles, behavior monitoring, and learning support. In situations where smaller groups may be necessary, there is also the Alternative Teaching method. With this method, small groups (less than 50% of the class) of students work with one teacher while the majority of the class works with the other teacher. This method is useful for pre-teaching or re-teaching content, as well as for enrichment groups (Cook & Friend, 1995). The final version of co-teaching is referred to as Team Teaching. In this model both teachers give content input simultaneously to the class. This final method is the most difficult to implement according to the literature because of the high amount of co-planning, teacher-to-teacher confidence, and role comfort required (Aliakbari & Nejad, 2013; Pappamihiel, 2012).

Chang (2018) mentions that in recent years the One Teaching-One Assisting method and the Alternative Teacher method have been expanded upon and focused on a more specific teacher role. Specifically, the One-Teaching-One Assisting method has been broken down to emphasize the primary role of one of the teachers as either assisting or observing. In the case where a co-teacher is required to simply observe a class and for example collect data, then that would be considered One Teaching-One Observing. On the other hand, if the co-teacher was actively engaged in supporting student learning or classroom management in some way, then that would be considered One Teaching-One Assisting. The Alternative Teaching method has likewise been divided and made more specific or focused. It has been divided into a Supplemental Teaching method, which would be used for remediation purposes, and an Alternative Teaching method, which would be used to give instruction of the same content, but via alternate instructional approaches. The Alternative Teaching method could be considered useful for small group instruction of lower ability or higher ability students. The emphasis for the model would be on small groups of learners needing diversification in instructional methods or content input.

*Disagreement: Co-Planning and Assessment.* There appears to be disagreement in the literature over whether to include co-planning aspects within the general co-teaching model. Cook and Friend (1995) do not include co-planning in the definition of the model but do mention it as a useful practice. According to the authors, co-teaching is fixated on the issues related to instruction by two teachers in one classroom, whereas co-planning may or may not be a standard practice for a teaching team. They point out that the practice of teachers co-planning a single unit for different groups of students would not fall under the purview of co-teaching. However, according to several other authors co-planning is an integral part of the co-teaching process, with co-assessment considered a basic tenet on how to co-teach and achieve positive results (Brendle et al., 2017; Brinkmann & Twiford, 2012; Taşdemir & Yıldırım, 2017). This disagreement may be linked to the popularity of the One Teacher-One Observer model. As mentioned, this model requires little in terms of preparation on the part of the co-teacher, for it quickly and easily allows a teacher to identify their role and responsibilities in the class (i.e. lesson planning, assessment, instruction, and behavior management). Therefore, many teachers who find themselves in a

co-teaching situation with little to no notice or co-teaching experience may regard this model as a natural fallback method.

*Gaps in Literature.* The literature was limited and lacked definitive support for co-teachings effectiveness regarding student learning outcomes and engagement (Aliakbari & Nejad, 2013; Magiera et al., 2006). While certain subject or content areas, like reading and language arts, appear to be more receptive to co-teaching methods, several meta-analyses have shown mixed quantitative support for co-teaching (Aliakbari & Nejad, 2013; Pappamihiel, 2012). Additionally, much of the data has focused on the special-needs population, not the ESL/ELL population, multilingual students, or bilingual education settings. These gaps make sense as the co-teaching concept grew out of the desire to support and mainstream special needs children within the general student population (Pappamihiel, 2012). Hopefully, this study will add insight into co-teaching's impact on student engagement in the classroom, specifically multilingual student groups and bilingual educational settings.

**Methodology**

This study utilized a convergent mixed-methods research design. The mixed methods design incorporates both qualitative and quantitative data that are merged and analyzed to give stronger insight into a research question (Creswell & Creswell, 2018).  The researcher felt that to understand the impact of co-teaching on student engagement it was necessary to measure both teacher and student perspectives. So, for this study, the researcher chose a mixed methods design to compare qualitative teacher data, which was collected through teacher interviews, and quantitative student data, which was collected through a reflective Likert Scale. Data were collected concurrently in a convergent design to compare collected data from the different sources.

*Participants.* The study included 25 participants (24 students and 1 teacher) from a 3rd grade English class in a German-English bilingual school. The 24 students included 11 boys and 13 girls, ages seven to eight. As a group, there were several home languages (L1s), though German was the majority L1. Eighteen out of the 24 students had German as an L1, while only three of the students spoke English as an L1. One student had recently relocated to the country and had limited proficiency in both English and German. This student was receiving weekly support in the form of pull-out ESL and German as a Second Language (DaZ) lessons. Those pull-out lessons were staggered so that that the student only missed 50% of their regular English/German lessons.

Co-teaching was a cornerstone of the school identity, and most classes were co-taught as a rule. The school focused on the benefits of lower student to teacher ratios, best use of teacher strengths, variety of instructional methods, and behavioral support. The school also supported the idea of teacher L1 continuity to assist student language immersion. This meant that whichever language the teacher designated as their school-L1 (German or English) was the language they used in the classroom, regardless of their own bilingual/multilingual abilities. In English language classes both teachers would only speak

English, and in German language classes both teachers would only speak German, with exceptions being made only in cases of serious student misunderstanding or situational urgency. In other classes, like Math, Science or History, teacher pairs were purposely scheduled to allow one English L1 teacher and one German L1 teacher. This bilingual pairing was done to support the English and German language learners simultaneously. Because of the importance of co-teaching at the school, the student participants had experienced some type of regular co-teaching in their previous lessons. The primary teacher was purposefully selected because of his previous experience co-teaching, and specifically his teaching English with this grade level (3+ years). Though he had had no professional development training with co-teaching methods, he had had several years of experience co-teaching at this school. The site location was the group's regular English classroom, and no changes were made to the class schedule.

*Design.* The study design used a convergent mixed methods approach (Demir & Pismek, 2018). Both qualitative and quantitative data were collected simultaneously. It involved seven 80-minute classes. At the end of each lesson student feedback was collected via a Likert Scale and a short teacher interview was conducted. Teacher interviews, while qualitative in nature, were converted into a quantitative measure to make engagement comparisons more effective. After giving feedback on the class engagement, the primary teacher was asked to rate the overall class engagement on a 10-point scale. This measure was compared in side-by-side analysis with the students' reflective feedback.

*Materials.* Student engagement was measured with a Likert Scale. This scale was based on the School Engagement in Mathematics Scale by Rimm-Kaufman (2010). The instrument was comprised of 10 items (see Figure 1 and Appendix A), each rated on a 5-point scale. For each item, students were asked to rate their level of agreement (1=Absolutely no, 2= No, 3= Not sure, 4= Yes, and 5=Absolutely yes) with each statement about the English class from that day. Question items measured students' emotional, social, and cognitive engagement. Some example statements included: 'Today I worked as hard as I could.', 'I talked about the lesson with other kids in class.', and 'Today I was bored.' Question 10 was reverse scored, so that higher scores on each item would coincide with overall higher lesson engagement. This final question was also used to evaluate the reliability of student answers. For example, if students filled out the scale indicating they were very engaged in the lesson they would have to mark every box in the extreme right column (see Figure 1) except for Question ten. If they marked Question ten in a contradictory manner, then the reliability of their answers would be considered questionable. Though the instrument designed by Rimm-Kaufman (2010) has been shown to be an effective measurement of class engagement, its reliability could have been weakened because of the self-reporting nature of the questionnaire, the multilingual environment / English level abilities of the students, the young age of the students, and the short time allotted for data collection (Rimm-Kaufman & Leis, 2015). The researcher introduced the scale to the students and reviewed each item in detail before the study began. The question items were explained and reviewed at several points throughout the study period to ensure participant understanding.

| Mark the box that best applies to each statement | ABSOLUTELY NO. | No. | Not Sure. | Yeah. | YES! |
|---|---|---|---|---|---|
| 1.  Today I worked as hard as I could. | | | | | |
| 2. I wanted to understand today's lesson really well. | | | | | |
| 3. I tried to learn as much as I could. | | | | | |
| 4. I thought a lot about English today. | | | | | |
| 5. I talked about the lesson with other kids in class. | | | | | |
| 6. I helped other kids with their work when they didn't know what to do. | | | | | |
| 7. I shared ideas and materials with other kids in class. | | | | | |
| 8. I saw other students helping each other in class. | | | | | |
| 9. Class was fun today. | | | | | |
| 10. Today I was bored. | | | | | |

*Figure 1: Student Engagement Questionnaire*

*Data Collection.* The population under investigation met twice per week in two 80-minute blocks. The primary teacher conducted two classes alone, while the remaining lessons were planned and carried out utilizing the various co-teaching methods. To measure student engagement during lessons a Likert Scale questionnaire was given out and collected at the end of every class. Students were given five to ten minutes to complete the questionnaire. The items chosen for the questionnaire measured behavioral dimension of engagement. Students' responses provided feedback on their feelings of engagement, and their perception of the teaching methods.

In the 1st and 7th lesson the researcher acted solely as an observer during the lesson, while the content input and lesson tasks were managed by the primary teacher.  Though this could have been considered a co-teaching exercise utilizing the Teacher-Observer model, these two lessons were considered the non-co-teaching element for the study and served as the basis for further comparisons. In the remaining five lessons all five models of co-teaching were utilized based on the planning needs of the primary teacher (i.e. One Teaching-One Assisting/Teacher-Floater, Station Teaching, Parallel Teaching, Alternate Teaching, and Team Teaching). The primary teacher and co-teacher/researcher normally used two co-teaching methods during a single lesson. The initial input and wrap-up/plenary would be conducted with the One-Teaching/One Assisting method or Team Teaching method, while the main class task would be planned around and utilize one of the remaining three models (i.e. Station Teaching, Parallel Teaching, or Alternate Teaching). The Station Teaching model was used twice in situations of content introduction or reinforcement. The Parallel Teaching model was used once in an assessment/feedback situation requiring student presentations and group discussions. The Alternate Teaching

model was used twice when lower-ability (LA), middle-ability (MA), or higher-ability (HA) student groups needed content reinforcement or enhancement.

After each lesson, an interview was conducted, and the primary teacher would relate their impression of the class's overall engagement and discuss any engagement issues observed during the lesson. Later the primary teacher would propose a value score for the class's engagement that day. This score was rated on a ten-point scale with ten being 'Fully engaged' and zero being 'Not engaged at all'.

*Data Analysis.* In this study two criteria were examined and compared: the teacher's perceived level of class engagement and the students' self-reported level of engagement. Data were compiled and analyzed manually by comparing average student-reported class engagement to teacher perceived engagement. Data analysis of the qualitative data was done by transforming the interviews with the primary teacher into a quantitative value. After each interview, the primary teacher was prompted to assign a level of class engagement for the day's lesson based on a 10-point scale, with 10-points being 'Fully engaged'. Data was analyzed to identify a correlation between teacher and class perceived engagement scores, changes in engagement levels, and differences between non-co-taught lessons and co-taught lessons.

As the self-reporting nature of the student questionnaire opened the possibility for issues of reliability in student answers the researcher made several adjustments during the data analysis process. In situations where students reported full engagement or full disengagement for questions 1-9 on the questionnaire, but for which they marked the reverse scored question 10 improperly (i.e. it contradicted the previous nine items), the researcher remarked the final question to reflect the student's obvious positive or negative attitude. It was felt by the researcher that this issue was caused by students misunderstanding the question item because of language issues, a lack of concentration or engagement during the questionnaire process, or purposeful mismarking. Additionally, in situations where students skipped or left question items blank, the researcher marked those items with the neutral value of three. This was done to keep student scores consistent. This issue was most likely caused by a lapse in concentration by the student or was related to improper formatting of the questionnaire.

**Results**

The dependent variables in this study were the levels of teacher perceived engagement and student self-reported engagement during co-taught and non-co-taught English lessons. All data were transformed by basing them on a maximum scale value of 50-points. The Likert Scale had a maximum point value of 50, with a maximum of 5 points allocated for each of the ten questions. The teacher's interview-based data, while collected on a 10-point scale with 10 associated with 'Full engagement', was transformed by a factor of five to make reliable comparisons between the two sets of data. The researcher designed this strategy of transforming the data by a factor of five so that both data groups could be easily compared. Outliers were evident and likely resulted from over and under reporting on the part of the

students. The researcher removed the scores of two students from the second data set to compensate for regular outliers and to create a more balanced description of mean scores for class-reported engagement (Figure 3). It was found that Student 1 had regularly over reported their scores, while Student 2 had regularly under reported their scores, for both students had improperly marked the reverse scored question number 10 on multiple occasions.

In the after-class teacher interviews a common theme of transition periods became apparent. The class teacher judged the students' ability to move around the classroom effectively as an element of engagement, following teacher instructions, and staying on task. This inner class movement was certainly an important issue for the class, as a great deal of instructional time could be lost when breaking into groups, if using the Alternate Teaching model, or when moving between stations when using the Station Teaching model. The class teacher pointed out, and the researcher concurred, that several specific students were often identified during these transition periods as being focal points of off-task behavior and slowing or disrupting the class's transition between lesson stages. The teacher made a point to discuss the issue of transitioning between groups and moving around the classroom with the class on several occasions. The teacher linked the issue of effective movement within the room to the achievement of lessons goals and to students' own time management.
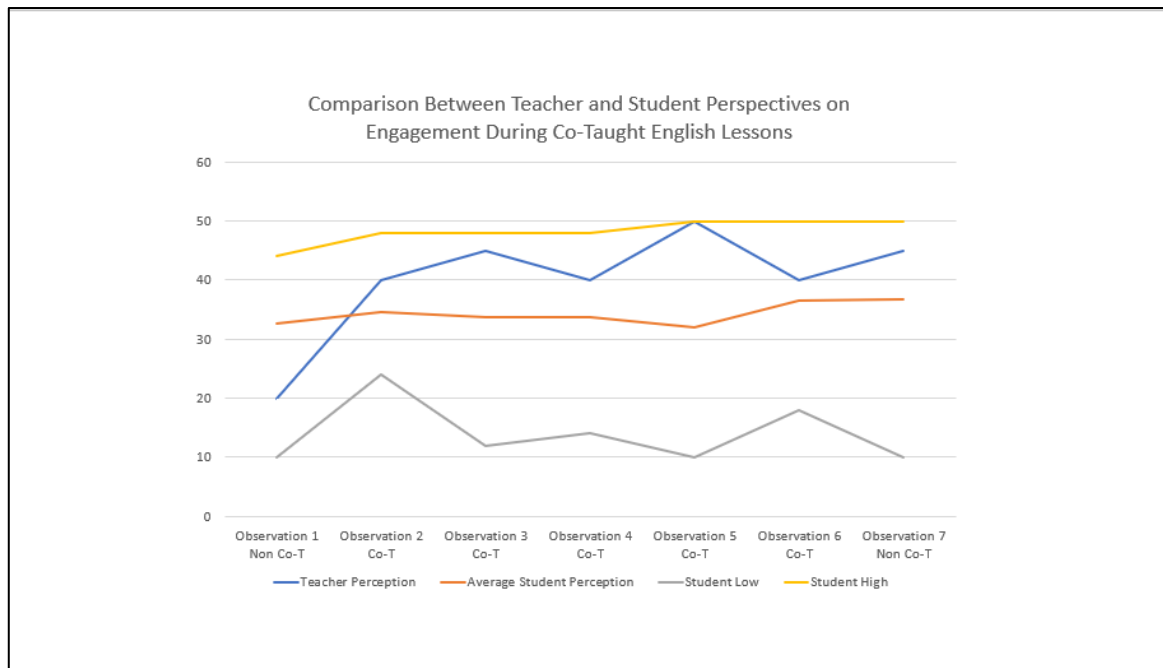


*Figure 2: Reported Teacher and Student Engagement Levels Including Outliers*

Figure two shows the initial data (Figure 2) containing the high and low outliers. In this presentation of data, the teacher and student-reported engagement levels only aligned once for Observation 5 and there only appeared to be a correlation between teacher-reported and student-reported engagement level increase or decrease between Observations 1 and 2, and minimally between Observations 6 and 7.  Overall there was a

wide disparity between teacher perception and student perception on engagement. Before accommodating for outliers there appeared to be no pattern to support an agreement between teacher and class perceived engagement scores, a positive correlation between engagement levels, or a pattern of perceived engagement or disengagement between non-co-taught lessons and co-taught classes.
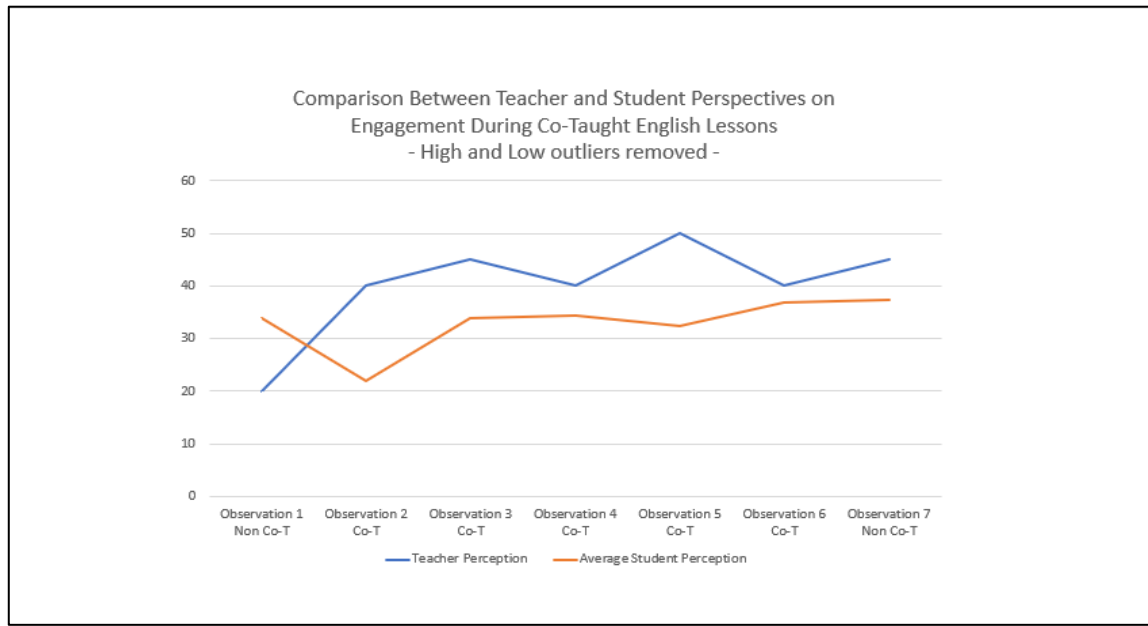


*Figure 3: Reported Teacher and Student Engagement Levels – Outliers Removed*

In Figure 3, the high and low outliers were removed.  With the outliers removed, student and teacher-reported scores did not align. The closest alignment of scores were found for Observation 6, which differed only slightly when compared to the data in Figure 2. However, it was apparent in the data that some of the differences between teacher-reported engagement scores and mean student scores increased. This was most obvious for Observation 2, which showed a drop in the student mean score of 40% compared to that of the initial data set. Furthermore, the change in perceived engagement between observations, which the researcher felt should have reflected a general agreement between teacher and class on the overall level of student engagement was only seen between Observation 1 and 2, and Observation 6 and 7. A lack of continuity between the reported engagement of non-co-taught and co-taught lessons was seen in the great differences between the initial and final non-co-taught lessons, and the irregular peaks and valleys evident among the co-taught lessons. While there appeared to be a positive correlation between engagement levels after the first lesson, this was minimal.

What was also interesting was the relative difference in teacher and student scores between co-taught and non-taught lesson. For six out of seven observations the teacher-reported scores ranged between 40 and 50 points. Coincidentally, for six out of seven observations the student-reported scores also fell within a ten-point range, specifically between 30 and 40 points. This may show a disagreement between teacher and students on

specific engagement levels, but simultaneously it may indicate a general agreement between the two groups.

## Discussion

This study found a lack of concrete support for co-teaching as a method of improving student engagement, as well as limited information related to the engagement of multilingual English learners. When data was adjusted for outliers, there was no agreement between teacher or students on engagement level scores, and the correlation of reported engagement levels between non-co-taught and co-taught lessons was minimal. This study can be linked to the work by Aliakbari and Nejad (2013) who cautioned on the use of co-teaching as popular model used by schools as a quick fix type method. Their study also showed a lack of student performance change based on these instructional methods.

Though there are many positive benefits from co-teaching, like lower student: teacher ratios, more instructional support, and more peer-to-peer learning, it has not been shown to directly improve student engagement or academic performance. The use of co-teaching as a buzz-word method or quick-fix solution to improve school perception or other issues could lead to instances of teachers being forced into co-teaching situations without having received any background on these instructional methods. This in turn could lead to negative perceptions of the methods that would further exacerbate unclear outcomes in engagement and academic progress. As mentioned previously, one issue that teachers must deal with, and which affects their impression of co-teaching is the balancing of roles among colleagues. Often the primary class teacher will view any co-teacher as an assistant, and naturally a co-teacher will most likely take on a supporting role if placed in another teacher's classroom. This is one reason why the One Teacher-One Assisting model is so popular. However, this negotiation of roles can be stressful for teachers, especially those that are not practiced or accustomed to co-teaching, and this can negatively impact their impression of co-teaching and their performance in the class.

## Limitations

Limitations in the study included a short timeline, student absences, regular outliers, participant age, participant language level, and questionnaire formatting. The study only included seven observations because of the timeline available. A longer-term study, with more scattered non-co-teaching observations, may have produced more generalizable data. Student absences limited the amount of data collected and may have skewed results. The researcher noticed several regular outliers in the data. One participant regularly gave the minimum score possible on the Likert Scale questionnaires. While this data was certainly interesting and attested to the participants' feelings about the lesson, it nonetheless skewed results considerably. This issue may have been a result of the participant's age or language ability, as several other instances of over and under scoring were apparent from different participants. Occasionally, this over and under scoring was obvious because of the reverse scored question in the scale, which the students mis-scored. Additionally, participants habitually left questions unanswered in the questionnaire. This was a larger

issue at the beginning of the study but was persistent throughout the data collection period. Both issues may have been caused by unclear formatting of the questionnaire or poor printing conditions. Finally, self-reflection is difficult for adults and so it must be assumed to be extra difficult for children. The student participants in this study were young and so their experience with self-reflection was limited and may have influenced their ability to properly evaluate their level of engagement.

**Implications**

Looking forward, there is a need for further investigations into the measurable benefits of co-teaching related to learning progress and lesson engagement. Additionally, more research is needed to see how these methods could best be used to support bilingual or multilingual learners. It may be very interesting to see a more quantitative study looking into the effects of co-teaching on student learning compared to non-co-teaching, such as examining the differences between a co-taught class and a non-co-taught class on achieving a learning objective or lesson module and assessment. Though much of the literature describes the benefits of co-teaching (e.g. best use of teacher strengths, improving student motivation, variety of instructional methods, and increase student learning) it must be remembered that the majority of these studies were originally conducted within a special education mindset and so there is a need to expand the literature on all fronts in order to include various student groups and needs (Magiera et al., 2006).

**Conclusion**

Though this study did not show a strong positive link between co-teaching and measurable student engagement in a multilingual classroom, it does not detract from the benefits of the method. With an additional perspective on the possible limitations of co-teaching, teachers can have a better understanding of what the method can and cannot do. Furthermore, teachers and administrators can use this study to discuss more frankly the purpose behind using co-teaching in the classroom and the expected outcomes when it is utilized. This study may show that, though co-teaching is not a guaranteed fix for improving student engagement, when compared to non-co-teaching it is still a successful model with many benefits for both teachers and students.

**About the Author**

**Kyle M. Daley** teaches English, Social Studies and Math in Berlin, Germany. Over the past ten years he has worked in Japan, China, and Germany at primary, secondary, and tertiary levels. He holds a M.Ed. in Curriculum and Instruction from the University of West Florida. His research interests include English Language Learners and Bi-lingual Education. Email: daley.kyl@gmail.com

## References

Aliakbari, M., & Nejad, A.M. (2013). On the effectiveness of team teaching in promoting learners' grammatical proficiency. *Canadian Journal of Education, 36*(3), 5-22. Retrieved from cje-rce.ca

Baca, L.M. (1990). Theory and practice in bilingual/cross cultural special education: Major issues and implications for research, practice, and policy. In *Proceedings of the First Research Symposium on Limited English Proficient Student Issues* (pp. 247–280). Washington, DC: U.S. Department of Education, Office of Bilingual Education and Minority Language Affairs. Retrieved from https://ncela.ed.gov

Brendle, J., Lock, R., & Piazza, K. (2017). A study of co-teaching identifying effective implementation strategies. *International Journal of Special Education, 32*(3), 538-550. Retrieved from https://www.internationalsped.com

Brinkmann, J. & Twiford, T. (2012). Voices from the field: Skill sets needed for effective collaboration and co-teaching. *International Journal of Educational Leadership Preparation, 7*(3), 1-13. Retrieved from http://www.ncpeapublications.org

Chang, S. H. (2018). Co-Teaching in Student Teaching of an Elementary Education Program, *Teacher Educators' Journal*, 11(1), 105-133. Retrieved from http://ateva.org

Cook, L. & Friend, M. (1995).  Co-teaching: guidelines for creating effective practices. *Focus on Exceptional Children*, *28*(3), 1-25. Retrieved from https://journals.ku.edu/FOEC/index

Creswell, J.W., & Creswell, J.D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Los Angeles: SAGE.

Magiera, K., Lawrence-Brown, D., Bloomquist, K., Foster, C., Figueroa, A., Glatz, K., Heppeler, D., & Rodriguez, P. (2006). On the road to more collaborative teaching: One school's experience. *TEACHING Exceptional Children Plus, 2*(5), 1-12. Retrieved from http://escholarship.bc.edu/education/tecplus

Maryland Department of Education. (2012). Effective co-teaching practices: A simple guide to co-teaching. Retrieved from https://www.anderson5.net/cms/lib/SC01001931/Centricity/Domain/3345/Co-Teaching%20Manual.pdf

Oh, K., Murawski, W., & Nussli, N. (2017). An international immersion into co-teaching: A wake-up call for teacher candidates in general and special education. *The Journal of Special Education Apprenticeship, 6*(1), 1-20. Retrieved from http://www.josea.info/

Pappamihiel, N. E. (2012). Benefits and challenges of co-teaching English learners in one elementary school in transition. *The Tapestry Journal, 4*(1), 1-13. Retrieved from http://tapestry.usf.edu/journal

Rimm-Kaufman, S. E. (2010). Student engagement in mathematics scale (SEMS). Unpublished measure, University of Virginia, Charlottesville, VA.

Rimm-Kaufman, S. E. & Leis, M. (2015). Using a partial credit model to evaluate the student engagement in mathematics scale. *Journal of Applied Measurement, 16*(3), 251-267. Retrieved from http://jampress.org/

Taşdemir, H., & Yıldırım, T. (2017). Collaborative teaching from English language instructors' perspectives. *Journal of Language and Linguistic Studies, 13*(2), 632-642. Retrieved from https://www.jlls.org/index.php/jlls

Yopp, RH., Ellis, M.W., Bonsangue, M.V., Duarte, T., & Meza, S. (2014). Piloting a co-teaching model for mathematics teacher preparation: Learning to teach together. *Issues in Teacher Education, 23*(1), 91-111. Retrieved from https://www.itejournal.org/

**Appendix A:  Student Engagement Questionnaire**

| Mark the box that best applies to each statement | ABSOLUTELY NO. | No. | Not Sure. | Yeah. | YES! |
|---|---|---|---|---|---|
| 1.  Today I worked as hard as I could. | | | | | |
| 2. I wanted to understand today's lesson really well. | | | | | |
| 3. I tried to learn as much as I could. | | | | | |
| 4. I thought a lot about English today. | | | | | |
| 5. I talked about the lesson with other kids in class. | | | | | |
| 6. I helped other kids with their work when they didn't know what to do. | | | | | |
| 7. I shared ideas and materials with other kids in class. | | | | | |
| 8. I saw other students helping each other in class. | | | | | |
| 9. Class was fun today. | | | | | |
| 10. Today I was bored. | | | | | |

- All items scored on a 5-point scale (1=Absolutely no, 2= No, 3= Not sure, 4= Yes, and 5=Absolutely yes)
- Item 10 reverse scored
- Note: Items based on Rimm-Kaufman (2010) Student Engagement in Mathematics Scale (SEMS).