Exploring the Efficacy of *The Word Within the Word* for Gifted and Typically Developing Students

Dr. Shelagh A. Gallagher
Engaged Education
Charlotte, NC 28211

Test measures used in this study are available by contacting Dr. Shelagh A. Gallagher, sgallagher5@carolina.rr.com.

Abstract

An exploratory study of the efficacy of *The Word Within the Word* tested students' ability to recognize, use, and recall vocabulary. Ten middle school teachers and their 493 students participated. Five teachers used *The Word Within the Word*, and five used traditional vocabulary materials. Students completed an out-of-level sentence completion test and a test of prompted vocabulary recall. Analysis of sentence completion data revealed significant differences with moderate effect sizes, favoring students in *The Word Within the Word* sixth- and seventh-grade classes. Analysis of prompted vocabulary recall data revealed significant differences with moderate to large effect sizes at all grade levels, favoring *The Word Within the Word* classes. Results suggest that in this case, both gifted and typically developing students in classrooms using *The Word Within the Word* were more skilled in vocabulary recognition, use, and recall than students in classrooms using traditional methods of vocabulary instruction.

Introduction

Possession of a substantial vocabulary is a hallmark of an educated mind, a cornerstone of eloquence, wisdom, and humor. Skilled choice of the appropriate word is at the heart of negotiation and collaboration. Vocabulary is a chief medium of conveying the story of ancient history and visions of future worlds. Ease with technical language distinguishes a novice from an expert. More pragmatically, a substantial vocabulary provides students with access to academic opportunities, including Talent Search programs, college admissions, and scholarships. For these and other reasons, gifted students should continue to build their vocabularies from kindergarten through twelfth grade.

Even so, little is known about the best methods to cultivate the vocabularies of advanced students. Research on vocabulary instruction has focused primarily on elementary grades, second language learners, or students who are academically at risk (Kame'enui & Maumann, 2012). A commonly recommended practice for these students is "incidental" vocabulary instruction, in which vocabulary is acquired as a result of reading quality literature (McCorquodale & Kirkland, 2006). It may seem reasonable to think that gifted students would learn advanced words as they read quality literature, but this is not guaranteed. Vocabulary acquisition through incidental reading is dependent on access to rigorous literature, and evidence suggests that literature assigned in middle and high school classrooms is substantially less challenging today than it was 20 years ago (Renaissance Learning, 2013). Even *bona fide* classics are not guaranteed to present students with an opportunity to develop their vocabularies. Thompson (2002) pointed out that the number of advanced words varies dramatically from classic to classic. By his count, *Of Mice and Men*, a novel with complex characters and worthy themes, has approximately 25 advanced vocabulary words, while *Tom Sawyer* has nearly 300. Moreover, incidental learning rarely goes beyond surface-level understanding of target words; for this reason, several experts have recommend that incidental vocabulary instruction is augmented with explicit instruction of words and their meanings (Biemiller, 2004; Lehr, Osborn, & Hiebert, 2004; Marzano, 2004; Nagy, 2005).

Direct instruction of vocabulary is usually designed to help students learn high-frequency words (Beck, McKeown, & Kocan, 2013). Recommended strategies to use during direct instruction include word play (Blachowicz & Fischer, 2012), graphic organizers and puzzles

(Lovitt, 1994), and morphological analysis (Carreker, 2005; Wilson, 2005). The primary objective of these strategies is to help students acquire vocabulary that will aid in reading comprehension, a limited goal that falls far short of adept language use or, for verbally gifted students, nuanced appreciation of words.

Vocabulary instruction for gifted students. Experts in gifted education have long recognized the need for advanced students to cultivate their verbal skills. As Michael noted, "Because language is man's chief means for receiving and transmitting knowledge, understanding of language is essential to progress. Gifted persons may be supposed, therefore, to need superior skill in the use of language and superior understanding of, and familiarity with, the media of language expression" (Michael, in Passow, 1996, p. 25). Given the importance of advanced vocabulary to academic success and professional expertise, vocabulary instruction has received surprisingly little attention in the gifted education literature. Even studies directly related to gifted students in language arts rarely include vocabulary outcomes (i.e., Feng, VanTassel-Baska, Quek, Bai, & O'Neill, 2005; Oh, Hailey, Azano, Callahan, & Moon, 2012; VanTassel-Baska, Johnson, Hughes, & Boyce, 1996).

Guidance on differentiating vocabulary instruction for gifted students has also been sparse. The only text on developing verbal ability in gifted students (Van Tassel-Baska, Johnson, & Boyce, 1996) has one chapter devoted to differentiated word play (Boyce, 1996) and one chapter on formal language study (Thompson, 1996). A graphic organizer called the Vocabulary Web, based on word etymology, has been recommended as a method of introducing students to etymological analysis (VanTassel-Baska & Little, 2011). These writings have provided some insight on how to vary vocabulary instruction, but none provides an organized hierarchy that could be the basis for comprehensive differentiation based on depth and complexity of vocabulary understanding.

Experts in vocabulary instruction have provided a scheme that might help establish a basis for systematic vocabulary differentiation. Nagy and Scott (2000) describe the different levels at which one can "know" a word. Surface word comprehension is at the bottom of their hierarchy, followed by the formation of schemas that describe relationships among words. Their next level of word knowledge is awareness of word connotations and other subtleties of word meaning. The highest level is word consciousness, the "awareness of and interest in words, their meanings and their power" (Lehr, Osborn, & Heibert, 2004, p. 16).

This hierarchy provides a helpful beginning but fails to provide a tier in which students come to know words with the level of sophistication needed to challenge gifted students, in which they "engage in language experiences at more complex, more abstract, more advanced, and more intense levels" (Passow, 1996, p. 30). Missing from this scheme, and from much of the literature on vocabulary instruction, is etymology. The study of word etymology shifts students' attention from components of word structure to the elements of word meaning, leading to a more insightful and exacting understanding of words. Bowers (2008) provided an example that distinguished morphology and etymology using the words *plea* and *please*. The words *plea* and *please* have different morphologies because each has a different base. In this case, each base is also a word. However, the suffix *-ant* can be added to the base *please* to create the word *pleasant*; the word *pleasant* cannot be formed by adding *-ant* to the base *plea*. Morphology helps students understand why the plural of *plea* is *pleas,* not *please,* and why the past tense of *plea* is *pled,* not *plead.*

If the study of *plea* and *please* ended with morphological analysis, students might conclude that the words are entirely different, but this is a misconception. Etymological analysis

takes students past surface structural differences and reveals a deeper conceptual similarity. *Plea* and *please* share common meaning through the Latin root *placere*, to please. The Latin root correctly unites *plea* and *please* with each other and with other words like *placebo*.

According to Thompson (2002), grounding vocabulary instruction in etymology has far-ranging benefits, including comfort with complex words, precise word choice, early acquisition of discipline-specific language, recognition of foreign language cognates, ease with spelling, and appreciation for language subtleties. Etymology provides a conceptual framework for understanding language, consistent with the advanced cognitive attributes observed in gifted adolescents (Gallagher, 2009). Adding etymology to the Nagy and Scott hierarchy in Table 1 introduces abstraction, depth, and complexity to the study of vocabulary, creating a bridge from word knowledge to a more substantial appreciation suggested by the term *word consciousness*. In the example in Table 1, a subtle but meaningful difference between the words *question* and *bequest* emerges only when their etymology is analyzed. The etymological analysis introduces new concepts of understanding that distinguish the words in ways that other levels of analysis do not.

The Word Within the Word. *The Word Within the Word* is a middle school vocabulary program that presents a different paradigm of instruction. *The Word Within the Word* uses etymology as its foundation, so students view vocabulary "…not as a set of lists of words but as a *system of thinking*, a way of building, analyzing, spelling, pronouncing, using, and choosing words" (Thompson, 2016, p. iii, emphasis added). The curriculum is based on the belief that students who are well-versed in vocabulary will be able to:

> …think intelligently about whether one word is more appropriate than another, more specific than another, more consonant than another with the rhythm and orchestration of the sentence, or more resonant in meaning than another. They can bring an array of criteria—cognitive, affective, and aesthetic—to critical thinking about word choice. (Thompson, 2002, p. 64)

The Word Within the Word is comprised of a series of vocabulary lessons that integrate the study of words with the study of Greek and Latin stems. Words and stems are selected to reveal deep-structure similarities among words that may appear different on the surface, allowing students to see their conceptual underpinnings. Students begin by memorizing stems and words so that they are immediately accessible, or automatized (Sternberg, 1997), much like memorizing multiplication tables. Once absorbed, students engage with the words and stems through activities that require a variety of different cognitive skills, as described in Table 2. Instead of focusing on high-frequency words, *The Word Within the Word* features words that regularly appear in classic literature. Throughout, students apply higher-order thinking to the study of language.

The Word Within the Word meets many of the curriculum modifications recommended for gifted students: the content is plentiful and advanced, and students learn to think about language systems and structures instead of discrete pieces of information (J. Gallagher et al., 1982; S. Gallagher, 2009; Van Tassel-Baska, & Little, 2011). Opportunities for creativity and originality are embedded in each lesson. Interdisciplinary thought is included through literary selections featuring words on the list. The multifaceted approach to word study is also consistent with recommendations for vocabulary instruction from the National Reading Panel (2000), which include the presentation of words in rich contexts and varying forms of vocabulary use.

Although *The Word Within the Word* enjoys increasing popularity, especially among teachers of the gifted, it has not been subjected to efficacy studies. The current study was designed as an initial exploration of achievement in classrooms using *The Word Within the Word* as compared to classrooms using traditional vocabulary instruction. Three research questions framed the study: (1) How do middle school students in classrooms using *The Word Within the Word* compare with middle school students in classrooms using traditional instructional methods on an out-of-level test of vocabulary recognition? (2) How do middle school students in classrooms using *The Word Within the Word* compare with middle school students in classrooms using traditional instructional methods on a test of prompted vocabulary recall? and (3) What differences in vocabulary recognition and use or prompted vocabulary recall are observed between gifted and typically developing students in classrooms using *The Word Within the Word* or traditional instructional methods?

Method

Participants

Teachers. Ten teachers and 493 students from six middle schools in a southeastern urban school district participated in the study. All 10 teachers volunteered to participate in the study, and all were provided a materials stipend in return for their participation. Each of the 10 teachers reported integrating vocabulary instruction every week throughout the school year.

Five of the 10 teachers used *The Word Within the Word* (WWW) materials in their language arts classes. The WWW teachers had between eight and 31 years of teaching experience, and they reported using the materials for four to 15 years. The WWW teachers differed in the amount of professional development they received related to the materials. One teacher had never attended professional development specific to WWW, one had attended one professional development session, and one had attended two sessions. The remaining two teachers had been to more than four professional development sessions specific to the curricula. Three of the five teachers used the materials in multiple course sections, resulting in a total of 12 classrooms using WWW materials.

The remaining five teachers used traditional instructional methods (TIMs) to teach vocabulary; four of the five responded to the teacher background survey. The four who responded had between one and 22 years of teaching experience. Like the WWW teachers, the TIMs teachers reported integrating vocabulary into their instructional plans each week. The teachers used a variety of commonly used strategies to engage their students in vocabulary study. The most frequently mentioned strategies were using puzzles and word games, asking students to memorize word lists or complete graphic organizers, and using words in context through sentence completion or analogical reasoning exercises. The more experienced teachers stated that they had been using these strategies throughout their careers. Only one of the four TIMs teachers had attended professional development devoted to vocabulary instruction. Two of the five TIMs teachers who participated in the study taught multiple class sections, resulting in a total of 11 TIMs classrooms.

Students. A total of 493 middle school students participated in the study, including 87 sixth graders, 200 seventh graders, and 206 eighth graders. Two hundred sixty-one of the 493 students met state criteria for gifted programs, which included a combination of demonstrated advanced reasoning ability (cognitive abilities at the 96th national age percentile using nationally

recognized measures), demonstrated academic achievement (94th national percentile or higher on approved achievement tests), and/or demonstrated academic performance (GPA of 3.75 out of 4.00). The study included 38 gifted students in sixth grade, 117 students in seventh grade, and 106 students in eighth grade.

Most students were in heterogeneously grouped classes, assigned as a part of the schools' regular scheduling procedures. The proportion of gifted and typically developing students in each classroom varied. Teachers using *The Word Within the Word* tended to have more gifted students; however, this was not always the case: one teacher using *The Word Within the Word* had more typical than gifted students, and one TIMs teacher had more gifted than typical students. Distribution of gifted and typically developing students across grade levels and teachers is presented in Table 3.

Materials

Teacher survey. Teachers completed a survey that asked how they approached vocabulary instruction. Teachers using *The Word Within the Word* were asked which sections of the curriculum they used; teachers using traditional curriculum were asked to list the instructional strategies they used during vocabulary instruction. All teachers were asked whether they approached vocabulary study as a year-round activity or as a specific unit of study. As reported previously, all of the teachers reported weekly instruction in vocabulary.

SAT Critical Reading Vocabulary test. A 20-item multiple-choice test was constructed from practice items in an SAT preparation book published by the College Board (2009). Test items were selected based on difficulty level as designated by the College Board; four items were selected at each of five difficulty levels. Students generally do not take the SAT until high school, so this test was considered out-of-level for both gifted and typically developing students.

The sentence completion items were designed by the College Board to assess students' knowledge of word meaning and their understanding of how words fit within a sentence. Each item was comprised of a sentence that was missing one or two words. Students had to select which of four possible words or word pairs best completed the sentence. To answer the question correctly, students had to understand both the meaning of the words and whether the words fit within the context of the sentence. A cronbach's alpha calculated on the administration of this test yielded a value of .68.

Prompted Vocabulary Recall test. Multiple-choice tests assess a student's ability to recognize a correct answer from among given choice; they do not measure a student's ability to produce a correct word from memory. A 10-item test was created for this study to assess students' ability to produce advanced vocabulary when prompted. Each item on the test provided students with a definition and a single word part—either a prefix, root, or suffix. Students were asked to produce the word that both used the word part and met the definition. For example, the correct answer for the item that presented the prefix *melan* and the definition "sadness or depression of the spirits" was *melancholy*. Words for this test were selected based on their frequent presence on eighth-grade word lists for a variety of school districts. The cronbach's alpha value for this test was .68.

Procedure

All teachers gave the tests to their students within the same two-week timeframe in the spring semester. Teachers received the tests on the first day of the administration period and had no prior knowledge about the tests. The tests were administered in a single 45-minute class period as part of regular instruction. Test forms were coded to create study variables for Grade Level (sixth, seventh, and eighth), Ability (Gifted or Typical), and Curriculum (WWW or TIMs) prior to analysis.

Data Analysis

Data analyses were conducted in three phases. In the first phase, a univariate analysis was conducted on demographic and classroom variables to identify potential covariates that should be included in the main analysis. In the second phase, separate 2x2x3 ANOVAs were calculated for the Sentence Completion and Prompted Vocabulary Recall tests to assess possible differences in performance according to Curriculum, Ability, and Grade Level. Type III Sums of Squares were used when interpreting the data. Using Type III Sums of Squares is advised when cell sizes are unbalanced; it also has the advantage of providing the variation attributable to any given variable after adjusting for the effects of other variables and interactions (Maxwell & Delaney, 2004). This approach allows for interpretation of a main effect even in the presence of interactions, which is particularly important in analyses comparing grade levels in which interactions due to maturation and experience are expected but not necessarily important. In the third phase, post hoc tests were conducted where indicated in the ANOVA results. The Games-Howell formula for paired contrasts was used for post hoc comparisons to account for unequal variance and cell sizes.

Effect sizes for the ANOVA and for the pairwise comparisons were calculated using Cohen's *d*. Cohen (1988) suggested a convention of 0.2 = small effect, 0.5 = medium effect, and 0.8 = large effect for interpretation; however, guidelines on how to interpret the magnitude of Cohen's *d* vary. Recent recommendations for interpreting effect size emphasize the importance of establishing contextual benchmarks as opposed to general guidelines; for example, average effect size for annual growth in elementary school is different from middle school (Coe, 2002; Hill, Bloom, Black & Lipsey, 2008). Hill and colleagues report that the average annual gain in effect size on standardized reading and mathematics tests in sixth grade are .23 and .30, respectively. In a meta-analysis of 36 studies, these authors found that the average effect size of an intervention in middle school is .27 (Hill et al., 2008). This is similar to Lipsey et al.'s (2012) report that average annual growth for middle school reading ranges from .23 to .26. Together, these add weight to Slavin's (2009) suggestion that an effect size of .25 is educationally relevant.

Results

Results of the data analysis revealed significant differences between students in classrooms using *The Word Within the Word* and traditional instructional materials. Differences were observed in overall comparisons of WWW and TIMs classrooms, between ability levels, and across grade levels.

Univariate Analysis of Student Demographic Data

The analysis of student demographic data yielded no statistically significant differences by gender. While there were indications of differences by racial/ethnic group, the proportional

difference in representation across groups disallowed including the variable in the analysis.

Average test scores by teacher and class section are presented in Table 4. The considerable variability in class composition across grade, teacher, and class section disallowed analysis nesting students within class, teacher, and school. A cursory review of the average scores reveals considerable variability across class sections for any given teacher, regardless of the curriculum she or he used. Most teachers who included only one class section in the study had average scores comparable to average scores of teachers with multiple sections. The differences observed in average scores by ability and by curriculum trended in the direction of the formal data analysis.

Analysis of Variance

Sentence Completion test. The three-way ANOVA yielded statistically significant main effects for Curriculum, Ability, and Grade Level. Average scores for students in each group are presented in Table 5; ANOVA results are summarized in Table 6.

The three-way ANOVA yielded a main effect for Curriculum, $F(1,481) = 9.20$, $p < .01$, $d = 0.20$, such that students using WWW ($M = 9.19$, $sd = 3.24$) scored statistically significantly higher than students using TIMs ($M = 7.89$, $sd = 3.55$). However, a statistically significant two-way interaction was observed for Curriculum x Grade, $F(2, 480) = 4.90$, $p < 0.00$, $d = 0.20$, suggesting that statistically significant differences by curriculum type might not be present at all grade levels.

A statistically significant main effect was also observed for Ability: $F(1, 481) = 52.59$, $p < .001$, $d = 0.55$. Gifted students scored statistically significantly higher than typically developing students (Gifted $M = 10.26$, $sd = 3.17$, Typical $M = 6.80$, $sd = 2.72$). A statistically significant interaction effect between Grade x Ability moderated the main effect: $F(2, 480) = 4.63$, $p < .01$, $d = 0.20$.

A third main effect was observed for Grade Level: $F(2, 480) = 12.51$, $p < .001$, $d = 0.41$. The average score for sixth-grade students ($M = 7.20$, $sd = 3.10$) was statistically significantly lower than for seventh-grade students ($M = 8.87$, $sd = 3.20$) and eighth-grade students ($M = 9.00$, $sd = 3.63$). No statistically significant differences were observed between seventh- and eighth-grade students.

No interaction effect was observed for Curriculum x Ability: $F(1, 481) = 0.38$, $p > .05$, $d = 0.00$. The three-way interaction of Curriculum x Ability x Grade Level was not statistically significant: $F(2, 480) = 0.79$, $p > .05$, $d = 0.00$.

Grade-level differences: WWW v. TIMs. A statistically significant two-way interaction was observed for Curriculum x Grade Level. Table 7 contains summary statistics for the Games-Howell post hoc test of these differences. The post hoc analysis yielded statistically significant differences between sixth-grade students in classrooms using WWW versus TIMs (WWW $M = 7.81$, $sd = 3.15$; sixth-grade TIMs $M = 5.29$, $sd = 2.00$; 5% critical difference = 1.75, mean difference = 2.53, $p < .001$, $d = 0.96$). A statistically significant difference favoring students in WWW classrooms was also observed in seventh grade (WWW $M = 9.87$, $sd = 2.98$; seventh-grade TIMs $M = 7.11$, $sd = 2.82$; 5% critical difference = 1.22, mean difference = 2.76, $p < .05$, $d = 0.74$). There was no statistically significant difference in Sentence Completion test scores of WWW and TIMs students in eighth grade (WWW $M = 9.22$, $sd = 3.37$; TIMs $M = 8.84$, $sd = 3.82$; 5% critical difference = 1.45, mean difference = 0.39, $p > .05$, $d = 0.30$).

Grade-level differences: Gifted v. Typical. Post hoc analysis of the interaction of Grade x Ability yielded no statistically significant difference between sixth-grade gifted and typical

students on the Sentence Completion test. Statistically significant differences were found in seventh grade, favoring gifted students (Gifted M = 10.41, sd = 2.72; Typical M = 6.69, sd = 2.50; 5% critical difference = 1.08 mean difference = 3.72, p < 0.001, d = 0.95). Statistically significant differences were also observed between eighth-grade gifted and typical students, favoring gifted students (Gifted M = 10.84, sd = 3.32; Typical M = 7.06, sd = 2.85; 5% critical difference = 1.24, mean difference = 3.78, p < .0001, d = 0.30). Summary statistics for this analysis are presented in Table 8.

Prompted Vocabulary Recall test. Average scores on the Prompted Vocabulary Recall test, by Grade Level, Ability, and Curriculum, are presented in Table 5; ANOVA results are included in Table 9. The three-way ANOVA of results on the Prompted Vocabulary Recall test yielded statistically significant main effects and interaction effects. A statistically significant main effect was observed for Curriculum, F(1,481) = 51.43, p < 0.001, d = 0.55, such that students using WWW scored statistically significantly higher than students using TIMs (WWW M = 3.52, sd = 1.83; TIMs M = 1.67, sd = 1.58). A statistically significant main effect was also observed for Ability, F(1, 481) = 40.28, p < 0.001, d = 0.46, in which gifted students scored statistically significantly higher than typically developing students (Gifted M = 3.52, sd = 1.83; Typical M = 167, sd = 1.58). A statistically significant main effect was also observed for Grade Level: F(2, 480) = 16.35, p < 0.001, d = 0.41.

Statistically significant two-way interactions were observed for Curriculum x Grade, F(2, 480) = 3.22, p < .05, d = 0.20, and Grade x Ability, F(2, 480) = 3.20, p < .05, d = 0.20, but not for Curriculum x Ability: F(1, 481) = 1.05, p > .05, d = 0.20. The three-way interaction of Curriculum x Grade x Ability was not statistically significant: F(2, 480) = 23.32, p > .05, d = 0.06.

Grade-level differences: WWW v. TIMs. Table 10 contains the results of the Games-Howell post hoc analysis of the Curriculum x Grade interaction. The analysis revealed statistically significant differences between students using WWW and TIMs, favoring WWW students in sixth grade (WWW M = 2.56, sd = 1.61; TIMs M = 0.43, sd = 0.60; 5% critical difference = 0.70, mean difference = 2.13, p < .000, d = 1.75), seventh grade (WWW M = 3.23, sd = 1.77; TIMs M = 1.63, sd = 1.43; 5% critical difference = 0.67, mean difference = 1.60, p < 0.001, d = 0.99), and eighth grade (WWW M = 3.91, sd = 2.17; TIMs M = 2.15, sd = 1.72; 5% critical difference = 0.81, mean difference = 1.76, p < 0.001, d = 0.90).

Grade-level differences: Gifted v. Typical. Post hoc analysis of the interaction between Grade x Ability revealed statistically significant differences favoring gifted students in sixth grade (Gifted M = 2.68, sd = 1.73; Typical M = 1.55, sd = 1.51; 5% critical difference = 1.05, mean difference = 1.11, p < .05, d = .70), seventh grade (Gifted M = 3.46, sd = 1.64; Typical M = 1.49, sd = 1.41; 5% critical difference = 0.62, mean difference = 2.00, p < .0.001, d = 1.29), and eighth grade (Gifted M = 3.89, sd = 1.97; Typical = 1.88, sd = 1.73; 5% critical difference = 0.71, mean difference = 2.11, d = 1.08). These results are presented in Table 11.

Discussion

The current exploratory study was designed to address three research questions: (1) How do middle school students in classrooms using *The Word Within the Word* compare with middle school students in classrooms using traditional instructional methods on an out-of-level test of vocabulary recognition knowledge? (2) How do middle school students in classrooms using *The*

Word Within the Word compare with middle school students in classrooms using traditional instructional methods on a test of prompted vocabulary recall? and (3) What differences in vocabulary recognition and use or prompted vocabulary recall are observed between gifted and typically developing students in classrooms using *The Word Within the Word* or traditional instructional methods? Analysis of the main effects and interactions in data suggest that students in classes using WWW were somewhat more skilled in word recognition and use and substantially more skilled in prompted vocabulary recall than students in classes using TIMs. Findings for the Sentence Completion test suggested that Ability had a significant impact on performance overall, but Curriculum had a significant impact in sixth and seventh grades. Findings on the Prompted Vocabulary Recall test revealed significant differences, with large effect sizes favoring students in WWW classes at each grade level. Achievement in WWW classrooms tended to be higher than achievement in TIMs classrooms regardless of student ability level, although gifted and typically developing students achieved at different levels.

Sentence Completion Test

The ANOVA of student responses on the Sentence Completion test yielded a significant effect for type of Curriculum, although the overall effect size was both small and smaller relative to the effect sizes for Grade Level and Ability. Grade Level, a variable that represents a combination of maturation and increased knowledge over time, had a smaller effect size than Ability. On the surface, this finding seems to suggest that student ability was the stronger determinant of performance on sentence completion multiple-choice test items found on the SAT Critical Reading test.

A different picture emerges when investigating the Curriculum x Grade Level interaction. Sixth- and seventh-grade students in WWW classes performed statistically significantly better than students in TIMs classes on the test of vocabulary recognition and use, with large effect sizes. In eighth grade, the difference in average scores, although favoring WWW students, was not statistically significant; this is likely the reason that there was no main effect for Grade Level in the analysis. However, even in eighth grade, the effect size measuring the magnitude of the difference between WWW and TIMs students exceeded Cohen's threshold for a small effect and matched the effect size of other educational interventions in middle school (Hill, Bloom, Black, & Lipsey, 2008).

Analysis involving grade and ability comparisons frequently result in statistically significant yet largely unimportant results. For example, one would expect gifted eighth-grade students to score significantly better than typically developing sixth-grade students on an achievement measure. In this case, non-significant findings in the cross-grade analysis were noteworthy. For instance, sixth-grade students in WWW classrooms scored higher than seventh-grade TIMs students, and seventh-grade WWW students scored higher than TIMs students in eighth grade—that is, the differences were not statistically significant where one would usually expect to see differences favoring students in higher grades. In this case, students in WWW classes performed up to a grade level beyond their typically developing age-mates on the out-of-level measure of word recognition and use.

Prompted Vocabulary Recall Test

As previously mentioned, multiple-choice questions are partially an assessment of word recognition; effective vocabulary instruction should also result in better word recall. The Prompted Vocabulary Recall test was designed for this study to determine whether students

could spontaneously recall a word when provided relevant prompts.

Results of the ANOVA on the Prompted Vocabulary Recall test were more straightforward than results of the Sentence Completion test. Scores of students in WWW classes were statistically significantly higher than those of students in TIMs classes on the Prompted Vocabulary Recall test overall and at each grade level. Cohen's *d* for the main effect of Curriculum was moderate, and it was larger than the effect size for Grade Level and Ability. Measures of Cohen's *d* comparing WWW and TIMs classrooms exceeded the threshold for a large effect at each grade level. Together, these findings suggest that the variable Curriculum had a greater influence than either Ability or Grade Level on the Prompted Vocabulary Recall test.

As with the SAT Sentence Completion test, analysis of the interaction effects on the Prompted Vocabulary Recall test produced some results that were noteworthy because of the absence of statistical significance. For instance, sixth-grade WWW students and eighth-grade TIMs students had statistically similar scores. This non-significant finding suggests superior achievement for sixth-graders in WWW classrooms, especially when contrasted with sixth-graders in TIMs classrooms, whose scores were significantly lower than those of eighth-grade students. When combined, these results suggest that students in classrooms using WWW could recall vocabulary better when prompted than TIMs students at the same grade level, and sometimes students in WWW classes remembered vocabulary with the same accuracy as TIMs students several grades higher.

The magnitude of the difference between WWW and TIMs students on the Prompted Vocabulary Recall test was larger and more consistent than on the SAT Sentence Completion test. This finding is somewhat counterintuitive: students who are better at recalling vocabulary might naturally be expected also to be better at recognizing vocabulary. Explaining this seeming anomaly is beyond the scope of this study, but several possibilities are worth exploring. First, the Prompted Vocabulary Recall test may have been easier than the SAT Sentence Completion test, since the Prompted Vocabulary Recall test was designed at grade level, and the SAT Sentence Completion test was out-of-level. Second, students in WWW classes may have benefitted more from the cue provided by the word stem included with each test item—that is, the test may have matched the format of the WWW materials. Third, results of the Sentence Completion test may have been influenced by a training effect, the result of the increasing amount of classroom time devoted to standardized test preparation. Gifted students, in particular, may have benefited from a practice effect if they had already completed the SAT to qualify for a Talent Search program.

Although *The Word Within the Word* was originally designed for gifted students, these results suggest that all students benefit from being in classrooms that use rigorous, cohesive curriculum materials. They also indicate that curriculum will not erase differences in ability. Typically developing students in WWW classes scored significantly higher than typically developing students in TIMs classes, but they did not routinely score higher than their gifted classmates taught using *The Word Within the Word*.

Differences by Ability Level

Ability had statistically significant main effects, with small to modest effect sizes in both ANOVA models. The interaction between Ability and Curriculum was not statistically significant, suggesting that each variable had an independent impact on performance on the two study measures. Interactions between Ability and Grade Level are explained to some extent by

statistically significant differences in comparisons of typical sixth-grade students and gifted eighth-grade students.

Grade-specific comparisons also indicated that Ability had a large effect on test performance that was independent of the Curriculum variable. Regardless of the form of curriculum used, gifted students scored higher on the study assessments than typically developing students. The magnitude of the effect size by Ability was larger than the effect size for grade-to-grade maturational or learning differences on both measures.

Limitations and Future Research

In the current exploratory study, both gifted and typically developing students in classrooms using WWW were superior to students of similar ability in classrooms receiving traditional instruction on measures of word recognition, use, and recall. What remains unanswered is why—a question that is hard to answer because of some structural limitations of this research. Chief among these is the lack a direct statistical control for prior learning, which would be especially useful in accounting for achievement differences between gifted and typically developing students and also in interpreting the results for teacher "I," whose students underperformed compared to other classrooms. Indeed, the absence of a pre-test or similar control for prior achievement makes it impossible to make an empirically-based inference that the patterns observed in the results of this study were due to the difference in materials used in WWW and TIMs classrooms.

This research does contain indicators that it would be worthwhile taking a closer look at the possible impact of *The Word Within the Word* on achievement. The assignment of students to heterogeneous classrooms within schools likely randomized the impact of prior learning in this study to some extent, as does the fact that gifted and typical students within classrooms had six months of shared instruction prior to the study. Student ability was accounted for, and the Curriculum variable had an effect independent of the Ability variable. General ability and achievement tend to share substantial covariance (Lubinski, 2000). Together, these trends suggest that the curriculum effect found in this study was not spurious. Regardless, additional research using pre- and post-test design would provide additional clarity to the current findings.

Differences in classroom composition also created an empirical limitation in the study, although arguably one that represents the real world of classroom assignment relatively well. Although statistical methods were used to control for differences in group sizes, some groups, particularly gifted sixth-grade TIMs students, were very small and must be interpreted with caution. Differences in distribution of gifted and typical students across classrooms also made analysis by teacher—and thus also by teaching experience or extent of professional development—difficult. In this case, it is impossible to tease apart the gestalt created by teacher experience, teacher knowledge, and classroom materials. However, it was interesting to note that classroom averages for teachers who taught multiple sections varied considerably, even when using the same materials in different classes. Even so, the advantage of *The Word Within the Word* over traditional instruction materials was consistent across different classrooms and schools. Statistically significant effects were found for Curriculum independent of Grade Level, and within-grade effect sizes between WWW and TIMs were moderate to large. So while it remains possible that the WWW classrooms were affected by an intervening variable unrelated to curriculum materials that was absent from the TIMs classrooms, the general direction of the current findings is cautiously suggestive of a curriculum effect.

Ideally, this initial exploration will serve as a springboard to additional research addressing the issues discussed here and taking new directions—for instance, investigating the efficacy of *The Word Within the Word* at different grade levels, using different outcome measures, and testing for knowledge retention over time. Other potentially fruitful directions to pursue are to assess the impact of professional development and fidelity to the curriculum model, gauging the impact of *The Word Within the Word* on student appreciation of language, and comparisons of *The Word Within the Word* with other structured vocabulary curriculum packages.

Despite the limitations, results of this initial exploratory study indicate that, in general, students in this study who were in classrooms using *The Word Within the Word* achieved at higher levels on measures of vocabulary knowledge and use, and especially on measures of vocabulary recall, than students in classrooms using more traditional approaches to instruction. Although gifted and typically developing students in WWW classrooms did not achieve at the same level, each group tended to achieve at higher levels in the WWW environment.

References

The American Heritage Dictionary (5th ed.). (2012). Houghton Mifflin.

Beck, I., McKeown, M. G., & Kocan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford.

Biemiller, A. (2004). Teaching vocabulary in the primary grades: Vocabulary instruction needed. In J. F. Baumann & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 28-40). Guilford.

Blachowicz, C. L. Z., & Fischer, P. (2012). Keeping the "fun" in fundamental: Encouraging word awareness and incidental word learning in the classroom through word play. In J. F. Baumann & E. F. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 219-238). Guilford.

Bowers, P. (2008). *WordWorks Newsletter #33*. Retrieved from http://wordworkskingston.com/WordWorks/Newsletter_33_-_links_to_PYP_files/WW%20%20%2333%20with%20comments.pdf

Boyce, L. (1996). In the big inning was the word: Word play resources for developing verbal talent. In J. VanTassel-Baska, D. T. Johnson, & L. N. Boyce (Eds.), *Developing verbal talent: Ideas and strategies for teachers of elementary and middle school students* (pp. 259-272). Allyn & Bacon.

Carreker, S. (2005). Teaching reading: Accurate decoding and fluency. In J. R. Birsh (Ed.), *Multisensory teaching of basic language skills* (2nd ed., pp. 43-81). Paul H. Brookes.

Coe, R. (2002). *It's the effect size stupid: What effect size is and why it is important.* Paper presented at the British Educational Research Association annual meeting, Exeter, England, September 12-14. Retrieved from www.cem.org/attachments/ebe/ESguide.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

College Board. (2009). *The official SAT study guide* (2nd ed.). Author.

Feng, A. X., VanTassel-Baska, J., Quek, C., Bai, W., & O'Neill, B. (2005). A longitudinal assessment of gifted students' learning using the integrated curriculum model (ICM): Impacts and perceptions of the William and Mary language arts and science curriculum. *Roeper Review, 27*(2), 78-83.

Gallagher, J. J., Kaplan, S. K., Passow, A. H, Renzulli, J. S., Sato, I. S., Sisk, D., & Wickless, J. (1982). *Principles of differentiated instruction*. Unpublished report. National/State Leadership Training Institute on the Gifted and Talented, Ventura, CA.

Gallagher, S. A. (2009). Designed to fit: Educational implications of gifted adolescents' cognitive development. In F. Dixon (Ed.), *Programs and services for gifted secondary students* (pp. 3-20). Prufrock Press.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172-177.

Kame'enui, E. J., & Maumann, J. F. (2012). *Vocabulary instruction: Research to practice* (2nd ed.). Guilford.

Lehr, F., Osborn, J., & Hiebert, E. H. (2004). *A focus on vocabulary*. Pacific Resources for Education and Learning.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms.* (NCSER 2013-3000). National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

Lovitt, S. (1994). Strategies for adapting science textbooks for youth with learning disabilities. *Remedial and Special Education, 15*(2), 105-116.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences in human behavior: "Sinking shafts at a few critical points." *Annual Review of Psychology, 51*, 405-444.

Marzano, R. J. (2004). *Building background knowledge for academic achievement: Research on what works in schools.* Association for Supervision and Curriculum Development.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective*. Taylor & Francis Group.

McCorquodale, G. J., & Kirkland, L. (2006). Building oral language and vocabulary through the use of literature. *Childhood Education, 83*(1), 32-m.

Nagy, W. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert, & M. L. Kamil (Eds.), *Teaching and learning: Bringing research to practice* (pp. 27-44). Lawrence Erlbaum.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269-284). Lawrence Erlbaum.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* National Institute of Child Health and Human Development.

Oh, S., Hailey, E., Azano, A., Callahan, C., & Moon, T. (2012). *What works in gifted education: Documenting the model-based curriculum for gifted students.* Society for Research on Educational Effectiveness, online serial, January.

Passow, A. H. (1996). Talent identification and development in the language arts. In J. VanTassel-Baska, D. T. Johnson, & L. N. Boyce (Eds.), *Developing verbal talent: Ideas and strategies for teachers of elementary and middle school students* (pp. 23-33). Allyn & Bacon.

Renaissance Learning. (2013). *What kids are reading: The book reading habits of students in American schools.* Author. Retrieved from http://doc.renlearn.com/KMNet/R004101202GH426A.pdf

Slavin, R. E. (2009). IBM's Writing to Read: Is it right for reading? *Phi Delta Kappan, 72*(3), 214-216.

Sternberg, R. J. (1997). A triarchic view of giftedness: Theory and practice. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 43-53). Allyn & Bacon.

Thompson, M. C. (1996). Formal language study for gifted students. In J. VanTassel-Baska, D. T. Johnson, & L. N. Boyce (Eds.), *Developing verbal talent: Ideas and strategies for teachers of elementary and middle school students* (pp. 149-173). Allyn & Bacon.

Thompson, M. C. (2002). Vocabulary and grammar: Critical content for critical thinking. *Journal for Secondary Gifted Education, 8*(2), 60-66.

Thompson, M. C. (2016). *The word within the word, volume I, teacher manual*. Royal Fireworks Press.

VanTassel-Baska, J., Johnson, D. T., & Boyce, L. N. (Eds.) (1996). *Developing verbal talent: Ideas and strategies for teachers of elementary and middle school students*. Allyn & Bacon.

VanTassel-Baska, J., Johnson, D. T., Hughes, C., & Boyce, L. N. (1996). A study of language arts curriculum effectiveness with gifted learners. *Journal for the Education of the Gifted, 19*(4), 461-480.

VanTassel-Baska, J., & Little, C. (Eds.). (2011), *Content-based curriculum for gifted learners.* Prufrock Press.

Wilson, B. A. (2005). Instruction for older students struggling with reading. In J. R. Birsh (Ed.), *Multisensory teaching of basic language skills* (2nd ed., pp. 319-344). Paul H. Brookes.

Table 1
A Hierarchy of Word Knowledge

Level	Type of Knowledge	Example using the word *question*	Example using the word *bequest*
Surface word meaning	Knows the current definition of the word	An expression of inquiry that invites or calls for a reply*	The act of giving or leaving personal property by a will*
Word morphology	Knows how the word is formed from a base in combination with a suffix or prefix	quest + ion	be + quest
Word schemas and families	Understands that additional words have similar spelling features, and these sometimes suggest similar meaning	inquest request conquest	inquest request conquest
Word etymology	Understands the historic source of the word; recognizes that knowing the source leads to more precise understanding	The origin of *question* is the Latin *quaerere*, to ask, seek.	The origin of *bequest* is the Latin *be* (cause to be or provide) and the Old English *cwis*, saying.
Word consciousness	Appreciates the system of language; takes pleasure or pride in the understanding	Appreciation that a question is a query but a bequest is a pronouncement	

Derived from Nagy and Scott (2000) and Bowers (2008), www.realspellers.org/forums/orthography/10-comments/unsorted-comments/941-morphology-etymology.

* Definitions for both words are from the *American Heritage Dictionary* (2012).

Table 2
Activities and Materials in a* Word Within the Word *Lesson

Section	Description
List	Latin and Greek stems are presented, along with a definition of each stem and a selection of vocabulary words that contain the stem.
Stem Close-Up	A stem from the list is introduced with specific meanings and variations. A separate set of stem-specific words is presented with this description.
Sentences	This is a list of sentences, each one using a vocabulary word introduced in the stem list, allowing students to see the word (and hence the stem) used in context.
Ideas	These pages contain descriptions of the ideas evoked by selected words that have been introduced with the stem list. For example, when students study the word *autodidacts,* they are introduced to examples from *Robinson Crusoe, Frankenstein,* and Wile E. Coyote.
Analogies	The Analogies pages contain 10 analogies that students must solve. Each set of analogies questions is made of words that use the stems from the stem list. Many of the answers also use words that contain stems from the stem list, continually reinforcing the stems and their meanings.
Notes	The notes are in-depth explorations of words. Subsections include Micropoems, which reveal philosophical kernels at the heart of word stems, and Classic Words, which present words commonly found in classic literature.
Classic Words	The Classic Words pages contain five sentences from classic literature, each one missing a word. Students must choose the correct word from the possible answers, and each possibility is a word that contains a stem from the stem list.
Tests	Cumulative tests require students to define both key stems and words. For *sublime,* for example, students must provide the definition of *sub,* as well as of the word *sublime.*

Table 3
Distribution of Students Across Grades and Teachers

	WWW			TIMs	
	Gifted	Typical		Gifted	Typical
Grade Level					
6th	34	32		4	17
7th	98	29		19	54
8th	55	34		51	66
Teacher			Teacher		
A	49	23	G	5	16
C	30	34	H	38	33
D	52	0	I	4	17
E	34	32	J	13	33
F	22	6	B	14	38

Table 4

Mean and Standard Deviation on the SAT Sentence Completion Test and Prompted Vocabulary Recall Test for Gifted and Typically Developing Students with Teachers Using WWW or TIMs in Sixth, Seventh, and Eighth Grades

				SAT Sentence Completion Test						Prompted Vocabulary Recall Test					
				Gifted			Typical			Gifted			Typical		
Grade	Curriculum	Teacher	Section	Mean	sd	n	Mean	sd	n	Mean	sd	n	Mean	sd	n
6th	WWW	E	1	9.57	2.95	14	7.44	2.65	9	3.29	1.94	14	2.78	1.72	9
			2	8.43	4.50	7	7.00	3.11	14	2.57	1.51	7	1.71	1.14	14
			3	7.85	2.38	13	8.22	3.10	9	2.69	1.38	13	2.22	1.72	9
	TIMs	I	1	5.50	1.00	4	5.24	2.19	17	0.50	0.58	4	0.41	0.62	17
7th	WWW	A	1	11.17	3.05	18	8.00	2.00	6	2.56	1.46	18	1.67	1.37	6
			2	8.75	3.27	20	9.75	2.76	8	3.15	1.50	20	2.25	1.58	8
			3	9.20	3.29	10	7.22	2.22	9	2.80	1.75	10	1.78	1.64	9
		D	1	10.70	2.30	27	na	na	na	4.26	1.48	27	na	na	na
		F	1	11.09	2.45	22	8.33	3.61	6	4.36	1.71	22	2.33	1.21	6
	TIMs	B	1	9.00	na	1	5.25	2.34	12	1.00	na	1	0.67	1.07	12
			2	na	na	na	6.00	2.00	10	na	na	na	1.30	1.42	10
			3	9.33	2.60	9	7.33	1.73	9	2.89	0.78	9	1.80	1.54	9
			4	10.00	1.41	4	6.63	2.13	8	2.50	0.58	4	1.88	0.99	8
		G	1	11.20	1.92	5	6.47	1.96	16	3.20	1.64	5	0.88	1.15	16

8th	WWW	C	1	8.71	2.56	7	7.73	1.74	11	3.57	2.14	7	2.19	1.66	11
			2	10.31	2.66	13	6.79	2.49	14	4.31	2.18	13	2.79	1.42	14
			3	8.10	3.41	10	7.67	2.74	9	3.10	1.52	10	3.00	2.29	9
		D	1	12.67	2.80	25	na	na	na	5.76	1.48	25	na	na	na
	TIMs	H	1	12.19	2.66	16	9.67	1.58	9	3.50	1.41	16	3.22	1.20	9
			2	9.00	2.65	3	7.30	2.92	20	1.33	0.58	3	0.65	0.75	20
			3	12.11	3.12	19	10.00	1.83	4	3.00	1.60	19	1.25	1.26	4
		J	1	9.38	4.27	8	7.68	1.89	19	2.88	1.46	8	1.37	1.01	19
			2	10.00	1.87	5	6.57	2.24	14	2.80	0.84	5	0.93	0.73	14

Table 5

Mean Scores of Students in WWW and TIMs Classrooms on the SAT Sentence Completion Test and Prompted Vocabulary Recall Test by Grade (6th, 7th, 8th), Ability (Gifted, Typical), and Grade x Ability

	SAT Sentence Completion Test									Prompted Vocabulary Recall Test								
	WWW			TIMs			Total			WWW			TIMs			Total		
	n	*M*	*sd*	*n*	*M*	*sd*				n	*M*	*sd*	n	*M*	*sd*	n	*M*	*sd*
Total	282	9.19	3.24	211	7.89	3.55				282	3.52	1.83	211	1.67	1.58			
Grade Level																		
6th	66	7.82	3.15	21	5.29	2.00	87	7.21	3.10	66	2.56	1.61	21	0.43	0.60	87	2.05	1.70
7th	127	9.87	2.98	73	7.11	2.82	200	8.87	3.20	127	3.23	1.77	73	1.63	1.44	200	2.65	1.82
8th	89	9.22	3.37	117	8.84	3.82	206	9.00	3.62	89	3.91	2.17	117	2.15	1.72	206	2.91	2.11
Ability																		
Gifted	187	10.10	3.15	74	10.65	3.19	261	10.26	3.17	187	3.80	1.90	74	2.82	1.45	261	3.52	1.83
Typical	95	7.39	2.59	137	6.39	2.74	232	6.80	2.72	95	2.28	1.58	137	1.25	1.44	232	1.67	1.58
Grade x Ability																		
6th																		
Gifted	34	8.47	3.26	4	5.50	1.00	38	8.16	3.23	34	2.94	1.63	4	0.50	0.58	38	2.68	1.73
Typical	32	7.13	2.92	17	5.24	2.19	49	6.47	2.81	32	2.16	1.51	17	0.41	0.62	49	1.55	1.51
7th																		
Gifted	98	10.50	2.81	19	9.94	2.20	117	10.41	2.72	98	3.59	1.70	19	2.79	1.08	117	3.46	1.64
Typical	29	7.76	2.56	54	6.11	2.29	83	6.69	2.50	29	2.00	1.44	54	1.22	1.33	83	1.49	1.41
8th																		
Gifted	55	10.40	3.39	51	11.31	3.22	106	10.34	3.32	55	4.69	2.05	51	3.02	1.46	106	3.89	1.97
Typical	34	7.32	2.32	66	6.92	3.10	100	7.06	2.85	34	2.65	1.74	66	1.48	1.60	100	1.88	1.73

Table 6

Analysis of Variance of Student Scores on the SAT Sentence Completion Test by Curriculum (WWW, TIMs), Grade Level (6^{th}, 7^{th}, 8^{th}), and Ability (Gifted, Typical)

Source	Type III SS	*df*	*F*	*p*	*d*
Curriculum	75.48	1	9.20	.003	0.20
Grade Level	205.34	2	12.51	.000	0.41
Ability	431.66	1	52.59	.000	0.55
Curriculum * Grade Level	80.37	2	4.90	.008	0.20
Curriculum * Ability	3.10	1	0.38	.539	0.00
Grade Level * Ability	76.06	2	4.63	.010	0.20
Curriculum * Grade Level * Ability	12.94	2	0.79	.455	0.00
Model	1838.45	11	20.36	.000	1.37
Error	3948.36	481			
Total	5786.81	492			

Table 7
Games-Howell Post Hoc Analysis of Grade Level (6th, 7th, 8th) x Curriculum (WWW, TIMs) for the SAT Sentence Completion Test

(I) Curriculum	(J) Curriculum	5% Critical Difference	Mean Difference (I-J)	*p* (2-tailed)	Cohen's *d*
6th WWW	6th Traditional	1.74778	2.53247	.001	0.96
	7th Word	1.35978	-2.05583	.001	0.76
	7th Traditional	1.47503	0.70859	.732	0.24
	8th Word	1.52731	-1.40654	.082	0.43
	8th Traditional	1.52069	-1.01943	.376	0.29
6th TIMs	6th Word	1.74778	-2.53247	.001	0.96
	7th Word	1.55055	-4.5883	.001	1.80
	7th Traditional	1.64711	-1.82387	.020	0.74
	8th Word	1.69092	-3.939	.001	1.42
	8th Traditional	1.68396	-3.55189	.001	1.16
7th WWW	6th Word	1.35978	2.05583	.001	0.76
	6th Traditional	1.55055	4.5883	.001	1.80
	7th Traditional	1.22371	2.76443	.001	0.95
	8th Word	1.28626	0.6493	.688	0.20
	8th Traditional	1.27838	1.03641	.175	0.30
7th TIMs	6th Word	1.47503	-0.70859	.732	0.24
	6th Traditional	1.64711	1.82387	.020	0.74
	7th Word	1.22371	-2.76443	.001	0.95
	8th Word	1.40753	-2.11513	.001	0.69
	8th Traditional	1.40034	-1.72802	.005	0.52
8th WWW	6th Word	1.52731	1.40654	.082	0.43
	6th Traditional	1.69092	3.939	.001	1.42
	7th Word	1.28626	-0.6493	.688	0.20
	7th Traditional	1.40753	2.11513	.001	0.69
	8th Traditional	1.45531	0.38711	.972	0.30
8th TIMs	6th Word	1.52069	1.01943	.376	0.29
	6th Traditional	1.68396	3.55189	.001	1.16
	7th Word	1.27838	-1.03641	.175	0.30
	7th Traditional	1.40034	1.72802	.005	0.52
	8th Word	1.45531	-0.38711	.972	0.30

Table 8
Games-Howell Post Hoc Analysis of Ability (Gifted, Typical) x Grade Level (6th, 7th, 8th) Interaction on the SAT Sentence Completion Test

(I) Ability	(J) Ability	5% Critical Difference	Mean Difference (I-J)	*p* (2-tailed)	Cohen's *d*
6th Gifted	6th Typical	1.9594	1.68851	.122	0.56
	7th Gifted	1.73795	-2.25236	.004	0.75
	7th Typical	1.76589	1.47115	.146	0.51
	8th Gifted	1.83039	-2.68173	.001	0.67
	8th Typical	1.77952	1.09789	.450	0.36
6th Typical	6th Gifted	1.9594	-1.68851	.122	0.56
	7th Gifted	1.40021	-3.94087	.001	1.42
	7th Typical	1.43617	-0.21736	.998	0.08
	8th Gifted	1.51711	-4.37023	.001	1.26
	8th Typical	1.45332	-0.59061	.837	0.21
7th Gifted	6th Gifted	1.73795	2.25236	.004	0.75
	6th Typical	1.40021	3.94087	.001	1.42
	7th Typical	1.07744	3.72351	.001	1.42
	8th Gifted	1.18517	-0.42937	.901	0.02
	8th Typical	1.10121	3.35026	.001	1.20
7th Typical	6th Gifted	1.76589	-1.47115	.146	0.51
	6th Typical	1.43617	0.21736	.998	0.08
	7th Gifted	1.07744	-3.72351	.001	1.42
	8th Gifted	1.22732	-4.15288	.001	1.24
	8th Typical	1.14645	-0.37325	.935	0.14
8th Gifted	6th Gifted	1.83039	2.68173	.001	0.67
	6th Typical	1.51711	4.37023	.001	1.26
	7th Gifted	1.18517	0.42937	.901	0.02
	7th Typical	1.22732	4.15288	.001	1.24
	8th Typical	1.24824	3.77962	.001	1.06
8th Typical	6th Gifted	1.77952	-1.09789	.450	0.36
	6th Typical	1.45332	0.59061	.837	0.21
	7th Gifted	1.10121	-3.35026	.001	1.20
	7th Typical	1.14645	0.37325	.935	0.14
	8th Gifted	1.24824	-3.77962	.001	1.06

Table 9

Analysis of Variance of Student Scores on Prompted Vocabulary Recall Test by Curriculum (WWW, TIMs), Grade Level (6^{th}, 7^{th}, 8^{th}), and Ability (Gifted, Typical)

Source	Type III SS	*df*	*F*	*p*	*d*
Curriculum	130.27	1	51.43	.001	0.55
Grade Level	82.85	2	16.35	.001	0.41
Ability	102.04	1	40.28	.001	0.46
Curriculum * Grade Level	16.33	2	3.22	.04	0.20
Curriculum * Ability	2.67	1	1.05	.31	0.20
Grade Level * Ability	16.23	2	3.20	.04	0.20
Curriculum * Grade Level * Ability	1.55	2	0.31	.74	0.06
Model	649.63	11	23.32	.001	1.47
Error	1218.36	481			
Total	1867.99	492			

Table 10
Games-Howell Paired Comparisons Grade Level (6th, 7th, 8th) x Curriculum (WWW, TIMs) for Prompted Vocabulary Recall Test

(I) Curriculum	(J) Curriculum	5% Critical Difference	Mean Difference (I-J)	*p* (2-tailed)	Cohen's *d*
6th WWW	6th Traditional	0.70108	2.13203	.001	1.75
	7th Word	0.73218	-0.66774	.087	0.40
	7th Traditional	0.75312	0.93047	.005	0.61
	8th Word	0.87994	-1.34951	.001	0.71
	8th Traditional	0.73502	0.40676	.596	0.25
6th TIMs	6th Word	0.70108	-2.13203	.001	1.75
	7th Word	0.60244	-2.79978	.001	2.12
	7th Traditional	0.63010	-1.20157	.001	1.09
	8th Word	0.77897	-3.48154	.001	2.19
	8th Traditional	0.60612	-1.72527	.001	1.34
7th WWW	6th Word	0.73218	0.66774	.087	0.40
	6th Typical	0.60244	2.79978	.001	2.12
	7th Traditional	0.66697	1.59821	.001	0.99
	8th Word	0.80743	-0.68177	.141	0.34
	8th Traditional	0.64646	1.0745	.001	0.62
7th TIMs	6th Word	0.75312	-0.93047	.005	0.61
	6th Typical	0.63010	1.20157	.001	1.09
	7th Word	0.66697	-1.59821	.001	0.99
	8th Word	0.82647	-2.27998	.001	1.24
	8th Traditional	0.67009	-0.52371	.209	0.34
8th WWW	6th Word	0.87994	1.34951	.001	0.71
	6th Typical	0.77897	3.48154	.001	2.19
	7th Word	0.80743	0.68177	.141	0.34
	7th Traditional	0.82647	2.27998	.001	1.24
	8th Traditional	0.81001	1.75627	.001	0.90
8th TIMs	6th Word	0.73502	-0.40676	.596	0.25
	6th Typical	0.60612	1.72527	.001	1.34
	7th Word	0.64646	-1.0745	.001	0.62
	7th Traditional	0.67009	0.52371	.209	0.34
	8th Word	0.81001	-1.75627	.001	0.90

Table 11
Games-Howell Post Hoc Analysis of Gifted (Gifted, Typical) x Grade Level (6th, 7th, 8th) Interaction for Prompted Vocabulary Test

(I) Ability	(J) Ability	5% Critical Difference	Mean Difference (I-J)	*p* (2-tailed)	Cohen's *d*
6th Gifted	6th Typical	1.0498	1.10687	.029	0.70
	7th Gifted	0.94902	-0.83776	.106	0.46
	7th Typical	0.95157	1.16392	.007	0.75
	8th Gifted	1.00354	-1.21003	.008	0.65
	8th Typical	0.95365	0.89789	.070	0.46
6th Typical	6th Gifted	1.0498	-1.10687	.029	0.70
	7th Gifted	0.77781	-1.94463	.001	1.21
	7th Typical	0.78116	0.05704	.000	0.04
	8th Gifted	0.84545	-2.3169	.001	1.33
	8th Typical	0.7837	-0.20898	.969	0.20
7th Gifted	6th Gifted	0.94902	0.83776	.106	0.46
	6th Typical	0.77781	1.94463	.001	1.21
	7th Typical	0.62214	2.00168	.001	1.29
	8th Gifted	0.70218	-0.37227	.641	0.24
	8th Typical	0.62576	1.73565	.001	0.94
7th Typical	6th Gifted	0.95157	-1.16392	.007	0.75
	6th Typical	0.78116	-0.05704	.000	0.04
	7th Gifted	0.62214	-2.00168	.001	1.29
	8th Gifted	0.70538	-2.37395	.001	1.40
	8th Typical	0.62935	-0.26602	.825	0.25
8th Gifted	6th Gifted	1.00354	1.21003	.008	0.65
	6th Typical	0.84545	2.31690	.001	1.33
	7th Gifted	0.70218	0.37227	.641	0.24
	7th Typical	0.70538	2.37395	.001	1.40
	8th Typical	0.70858	2.10792	.001	1.08
8th Typical	6th Gifted	0.95365	-0.89789	.070	0.46
	6th Typical	0.7837	0.20898	.969	0.20
	7th Gifted	0.62576	-1.73565	.001	0.94
	7th Typical	0.62935	0.26602	.825	0.25
	8th Gifted	0.70858	-2.10792	.001	1.08