Meta-Analytic Synthesis of Studies Conducted at Marzano Research Laboratory on Instructional Strategies

By

Mark W. Haystead
&
Dr. Robert J. Marzano

Marzano Research Laboratory
Englewood, CO

August, 2009

Table of Contents

Executive Summary

This report synthesizes a series of quasi-experimental studies conducted as action research projects regarding the extent to which the utilization of selected instructional strategies enhances the learning of students. Over 300 volunteer teachers conducted independent studies at 38 schools in 14 school districts between fall 2004 and spring 2009. The data used for analysis can be found in Marzano Research Laboratory's Meta-Analysis Database (see marzanoresearch.com).

The independent studies involved 7,872 students in the experimental groups and 6,415 students in the control groups. Participating teachers selected two groups of students both of which were being taught the same unit or set of related lessons. However, in one group (the "experimental" group) a specific instructional strategy was used (e.g., graphic organizers), whereas in the other group (the "control" group) the instructional strategy was not used. Because students could not be randomly assigned to experimental and control groups, all studies employed a quasi-experimental design, referred to as a pretest-posttest non-equivalent groups design. The pretest scores were used as a covariate to partially control for differing levels of background knowledge and skill.

The following questions were considered through a meta-analysis of the 329 independent studies:

1. What effect does the utilization of instructional strategies have on students' achievement regarding the subject matter content taught by their teachers?
2. Does the effect of instructional strategies differ between school levels?
3. Does the effect of instructional strategies differ from strategy to strategy?

The average effect size for all 329 independent studies was statistically significant ($p < .0001$). When corrected for attenuation, the percentile gain associated with the use of the instructional strategies is 16 ($\overline{ES} = .42$). This means that on the average, the strategies used in the independent studies represent a gain of 16 percentile points over what would be expected if teachers did not use the instructional strategies.

Introduction

This report synthesizes a series of quasi-experimental studies conducted as action research projects regarding the extent to which the utilization of selected instructional strategies enhances the learning of students. Over 300 volunteer teachers conducted independent studies at 38 schools in 14 school districts between fall 2004 and spring 2009. The data used for analysis can be found in Marzano Research Laboratory's Meta-Analysis Database (see marzanoresearch.com).

Action Research Projects

Participating teachers selected two groups of students both of which were being taught the same unit or set of related lessons. However, in one group (the "experimental" group) a specific instructional strategy was used (e.g., advance organizers), whereas in the other group (the "control" group) the instructional strategy was not used. Because students could not be randomly assigned to experimental and control groups, all studies employed a quasi-experimental design, referred to as a pretest-posttest non-equivalent groups design. These groups are considered to be non-equivalent, because it is unlikely that two intact groups would be as similar as would be the case if randomly assigned.

A pretest and posttest was administered to students in both groups. The pretest scores were used to statistically "adjust" the posttest scores using a technique referred to as analysis of covariance (ANCOVA). In basic terms, the adjustment translates the posttest scores into those that would be expected if students in both groups started with the same scores on the pretest. In effect, it is a way of controlling for students' differences in what they know about a topic prior to the beginning of instruction on the topic. ANCOVA is commonly used when random assignment is not possible (see Technical Note 1). Although ANCOVA was used to statistically equate students in terms of prior academic knowledge, arguments about causal relationships are not as strong as they would be when group members are assigned through a random lottery.

Again, teachers were instructed to teach a short unit on a topic of their choice to two groups of students—one experimental and one control. Instructional activities in both groups were to be as similar as possible except for the fact that the instructional strategy was used in one group only (i.e., the experimental group). Directions provided to teachers are reported in Appendix A.

The Use of Meta-Analysis

Meta-analytic techniques (see Hedges & Olkin, 1985; Lipsey & Wilson, 2001; Cooper, 2009) were used to aggregate the findings from the independent studies using the statistical software package Comprehensive Meta-Analysis (CMA, Version 2). In general, meta-analytic techniques are used when the results of independent studies on a common topic are combined. For example, assume 25 studies were conducted in various sites on the effects of a specific instructional technique on student achievement. The studies were different in terms of the subject areas that were addressed. Consequently, different assessments of student achievement were used to reflect the different subject areas. This is the classic scenario requiring the use of meta-analytic techniques—independent studies on a common topic (i.e., a common instructional technique) but with different dependent measures.

To combine studies that used different dependent measures, the results of each study are translated into an effect size. While there are many types of effect sizes, the one used in this meta-analysis is the standardized mean difference. In very general terms, a standardized mean difference is the difference in the average score of the control group and the experimental group stated in standard deviation units. Thus, an effect size of 1.00 would indicate that the average score in the experimental group is one standard deviation higher than the average score in the control group. Conversely, an effect size of -1.00 would indicate that the average score in the experimental group is one standard deviation lower than the average score in the control group.

The present meta-analysis is analogous to this situation. A common class of interventions was used in all experimental classes (i.e., use of selected instructional strategies), but the independent studies employed teacher designed assessments of student academic achievement across various grade levels and subject areas requiring different dependent measures.

Meta-analytic findings are typically reported in two ways, 1) findings based on the observed effect sizes from each independent study (see Appendix B), and 2) findings based on a correction for attenuation due to lack of reliability in the dependent measure (i.e., teacher designed assessments of student academic achievement). Technical Note 2 explains the method used to correct for attenuation and an interpretation of such corrections. Briefly though, when a dependent measure is not perfectly reliable it will tend to affect the strength of observed relationships between independent and dependent variables.

An independent variable is a factor which is assumed or hypothesized to have an effect on some outcome often referred to as the dependent variable. A dependent variable is an outcome believed to be influenced by one or more independent variables. For this meta-analysis of the independent studies, the dependent variable was students' knowledge of academic content addressed during a unit of instruction and the independent variable of interest was the use of the selected instructional strategy (e.g., feedback). It is always advisable to correct an effect size for attenuation (i.e., decrease in effect size) due to unreliability of the dependent measure (for a detailed discussion of attenuation see Hunter & Schmidt, 2004). In basic terms, every assessment is imprecise to some extent and this imprecision lowers the effect size. Throughout this report, observed and corrected effect sizes are displayed for comparison. When this is the case, the discussion of findings is limited to the corrected results only.

The Sample

Figure 1 displays the number of participating sites and independent studies by school level along with the number of students in experimental and control groups.

Figure 1. Number of Participating Sites and Independent Studies by School Level

School Level	# of Sites	N	Cn	En	Tn
Elementary School (Grades K-5)	19	55	1,040	1,041	2,081
Middle School (Grades 6-8)	8	64	1,527	2,710	4,237
High School (Grades 9-12)	11	210	3,848	4,121	7,969
Total	**38**	**329**	**6,415**	**7,872**	**14,287**

In all, this meta-analysis of the 329 independent studies involved 14,287 students. Of those students, 2,081 were at 19 sites that teach students at the elementary school level, 4,237 were at 8 sites that teach students at the middle school level, and 7,969 were at 11 sites that teach students at the high school level.

Data Analysis and Findings

As mentioned previously, in this meta-analysis one dependent variable was considered: students' knowledge of academic content addressed during a unit of instruction. The independent variable of interest was the experimental/control condition—whether students were exposed to an instructional strategy or not. Also of interest was the difference in potential effect of the utilization of instructional strategies at the elementary, middle, and high school levels.

Data from each independent study was first analyzed using the general linear model as employed by the statistical software package, SPSS (v17.0). One independent variable (experimental/ control condition) was entered into the equation using a fixed-effect model. (See Technical Note 3 for a discussion of fixed effects.) The dependent variable was the posttest scores with the pretest scores used as the covariate. Stated differently, a fixed-effects analysis of covariance (ANCOVA) was executed for each independent study. The ANCOVA findings were used to compute an effect size (i.e., standardized mean difference effect size) for each independent study (see Technical Note 4 for a discussion regarding the formula used to compute the effect size). CMA was then used to aggregate the findings from the independent studies using the observed and corrected effect sizes for the experimental/control condition (i.e., use of a selected instructional strategy).

Again, three questions were considered in this meta-analysis:

1. What effect does the utilization of instructional strategies have on students' achievement regarding the subject matter content taught by their teachers?
2. Does the effect of instructional strategies differ between school levels?
3. Does the effect of instructional strategies differ from strategy to strategy?

Findings for each question are discussed separately.

Question 1: What effect does the utilization of instructional strategies have on students' achievement regarding the subject matter content taught by their teachers?

Considered in isolation, most of the independent studies (see Appendix B) did not exhibit statistical significance. For an individual study to be considered statistically significant, the reported *p*-value should be less than .05 (see Murphy & Myors, 2004). According to this criterion, 90 of the 329 studies (or 27%) can be considered statistically significant. When the results of a set of studies are combined using meta-analytic techniques, the findings considered as a group might be statistically significant even though a number of the individual studies are not significant. Such is the case with the present set of studies. In fact this is quite common in educational research where many individual studies might be deemed non-significant simply because they do not have enough subjects in the experimental and control groups. However, when these studies are combined using meta-analytic techniques the aggregate finding is often highly significant (for a detailed discussion see Hedges & Olkin, 1985).

Figure 2 shows the overall average effect size for a meta-analysis of the 329 independent studies using a random-effects model of error (see Technical Note 5 for discussion of fixed- vs. random-effects meta-analysis). The column labeled "N" identifies the number of studies included in the group, the column labeled "$\overline{ES}$" reports the weighted average effect size for the studies, the column labeled "SE" contains the standard error for the reported weighted average effect size, the column labeled "95% CI" identifies the 95 percent confidence interval (lower limit and upper limit) for the reported weighted average effect size, the column labeled "Sig." reports the *p*-value for the reported weighted average effect size, the column labeled "% Gain" contains the percentile gain (or loss) associated with the reported weighted average effect size, and the column labeled "Fail-Safe N" identifies the number of missing studies that would be required to reduce the weighted average effect size to .01 using Orwin's formula (for a discussion of sampling bias and the fail-safe N, see Lipsey & Wilson, 2001, pp. 165-166).

Figure 2. Overall Random Effects for Instructional Strategies

	N	$\overline{ES}$	SE	95% CI		Sig. (2-tailed)	% Gain	Fail-Safe N
				LL	UL			
Overall	329	.36 (.42)	.03 (.04)	.30 (.35)	.43 (.50)	.000 (.000)	14 (16)	11,515 (13,489)

Note: Corrected findings are presented in parentheses.

When the results of the 329 independent studies are corrected for attenuation and combined, the overall effect size is .42 which is associated with a 16-percentile-point gain. This means that on the average, the instructional strategies used in the independent studies represent a gain of 16 percentile points over what would be expected if teachers did not use the instructional strategies (for a discussion of how effect sizes are combined and an overall significance level is computed see Lipsey & Wilson, 2001).

Consider the fail-safe N reported in parentheses, 13,489. This means that over 13,400 additional independent studies with an effect size of .00 would be needed to reduce the weighted average effect size to .01. The percentile gain associated with an effect size of .01 is 0 (i.e., no difference between groups).

The column labeled "95% CI" contains the 95 percent confidence interval for the reported weighted average effect size. Again, the effect size reported in Figure 2 is a weighted average of all the effect sizes from the 329 independent studies (see Appendix B). As such, it is considered an estimate of the true effect size of the experimental condition (i.e., use of instructional strategies). The 95 percent confidence interval includes the range of effect sizes in which one can be certain the true effect size falls. For example, consider the 95 percent confidence interval reported in parentheses, .35 to .50. This indicates a 95 percent certainty that the true effect size for the meta-analysis of the 329 independent studies is between the values of .35 and .50. When the confidence interval does not include .00, the weighted average effect size is considered to be statistically significant ($p < .05$). In other words, $\overline{ES} = .00$ would not be considered a reasonable assumption. In fact, the p-value associated with the reported effect size is less than .0001 indicating it is highly significant in laymen's terms. (For a detailed discussion of the meaning of statistical significance, see Harlow, Muliak, & Steiger, 1997.)

Another way to examine the general effect of the instructional strategies is to consider the distribution of effect sizes as shown in Figure 3.

Figure 3. Distribution of Effect Sizes

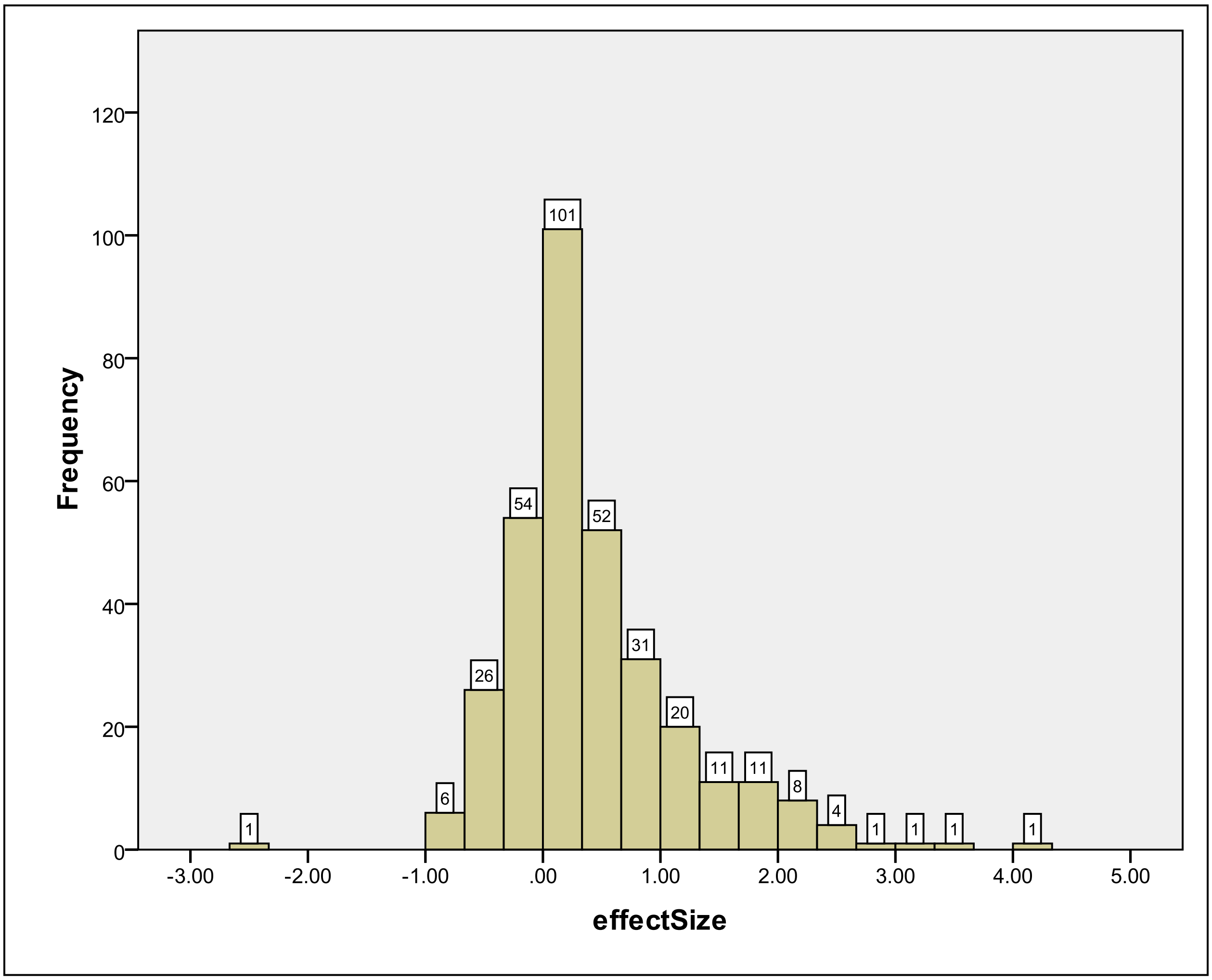


Figure 3 reports the distribution of "groups" of effect sizes across the 329 independent studies (see Appendix B). 87 studies exhibited a negative effect (see first through fourth columns), 184 studies exhibited an effect size between .00 and 1.00 (see fifth through seventh columns), 42 studies exhibited an effect size between 1.00 and 2.00 (see eighth through tenth columns), and so on. 242 out of 329 studies (or 74%) have a positive effect size.

Question 2: Does the effect of instructional strategies differ between school levels?

To address this question, a meta-analysis was employed using the school level for each independent action research study as a moderator variable. A moderator variable is a qualitative or quantitative factor that affects the direction and/or strength of the relation between the dependent and independent variables. The findings are reported in Figures 4 and 5.

Figure 4. Random Effects for School Level

School Level	N	$\overline{ES}$	SE	95% CI		Sig. (2-tailed)	% Gain
				LL	UL		
Elementary School (Grades K-5)	55	.65(.74)	.08(.09)	.48(.56)	.81(.93)	.000(.000)	24(27)
Middle School (Grades 6-8)	64	.29(.34)	.07(.08)	.15(.17)	.43(.50)	.000(.000)	11(13)
High School (Grades 9-12)	210	.31(.36)	.04(.05)	.23(.27)	.40(.46)	.000(.000)	12(14)

Note: See discussion of Figure 2 for a description of column headings. Corrected findings are presented in parentheses.

Figure 5. Homogeneity Analysis for School Level

Q	Sig. (2-tailed)	df
13.945 (14.255)	.001 (.001)	2

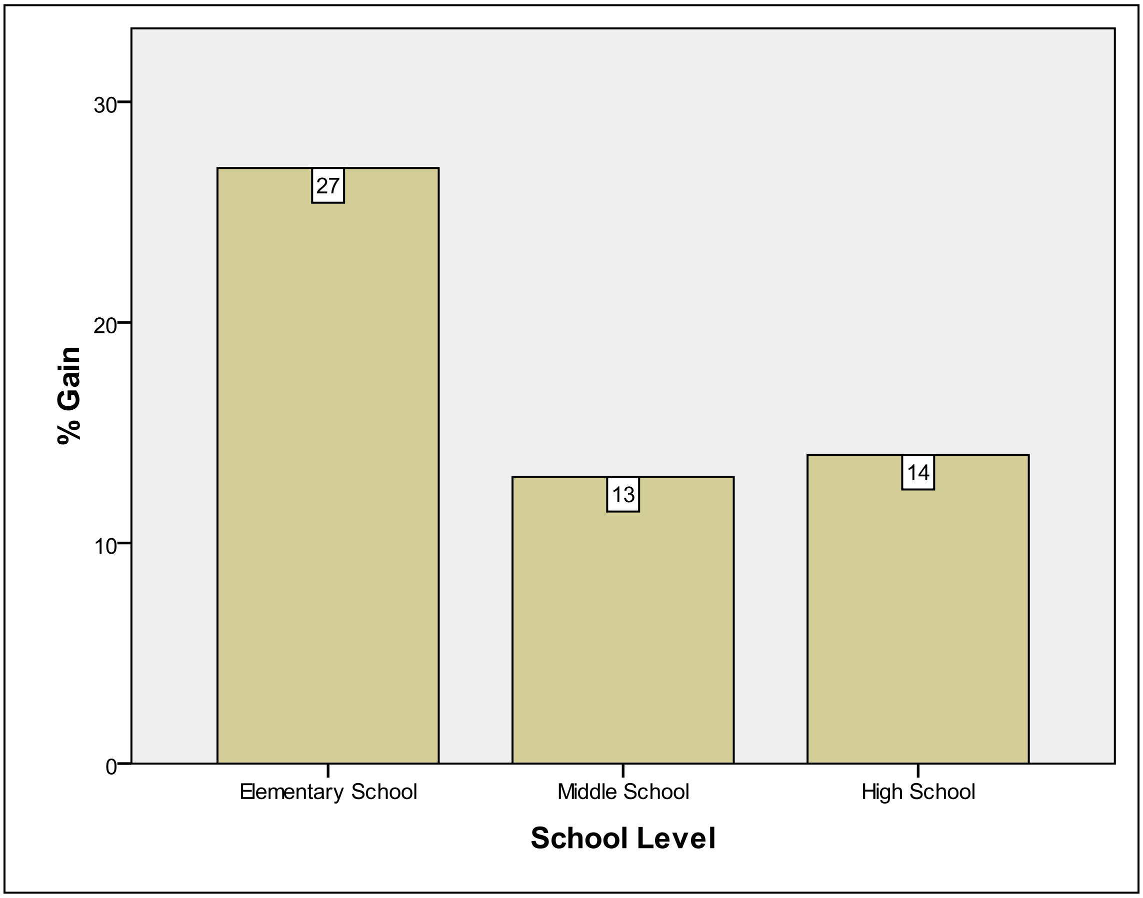
Note: Corrected findings are presented in parentheses.

Figure 4 shows the random effects for the elementary, middle, and high school levels. The weighted average effect size was statistically significant at the .0001 level ($p < .0001$) for elementary and high school levels and at the .001 level ($p < .001$) for middle school.

Figure 5 reports the results of the homogeneity analysis for the levels of the moderator variable—in this case school levels. A significant finding would indicate that the average effect sizes for the various levels of schooling most probably represent different populations. Stated differently, a significant *Q*-value would indicate that the means for the levels of schooling are significantly different. In this case, the *Q*-value was highly significant ($p < .001$).

Figure 6 graphically depicts the average percentile gains associated with each school level.

Figure 6. Percentile Gain for Random Effects for School Level (Corrected)



Question 3: Does the effect of instructional strategies differ from strategy to strategy?

To address this question, a meta-analysis of the 329 independent studies that utilized each of the 15 instructional strategies listed below was employed (for a discussion of the research and theory regarding some of these strategies, see Marzano, 2007). Only strategies that involved five or more studies were considered for this analysis. (For a complete listing of strategies found in the Meta-Analysis Database, see marzanoresearch.com.) Figures 7 and 8 present the findings for this analysis.

- Advance organizers – involves providing students with a preview of new content.
- Building vocabulary – involves use of a complete six step process to teaching vocabulary that includes: teacher explanation, student explanation, student graphic or pictographic representation, review using comparison activities, student discussion of vocabulary terms, and use of games. (For additional information on the six step process see Marzano, 2004, pp. 91-103.)
- Effort and recognition – involves reinforcing and tracking student effort and providing recognition for achievement.
- Feedback – involves providing students with information relative to how well they are doing regarding a specific assignment.
- Graphic organizers – involves providing a visual display of something being discussed or considered, e.g., using a Venn diagram to compare two items.
- Homework – involves providing students with opportunities to increase their understanding through assignments completed outside of class.
- Identifying similarities and differences – involves the identification of similarities and/or differences between two or more items being considered.
- Interactive games – involves use of academic content in game-like situations.
- Nonlinguistic representations – involves providing a representation of knowledge without words, e.g., a graphic representation or physical model.
- Note taking – involves recording information that is considered to be important.
- Practice – involves massed and distributed practice on a specific skill, strategy, or process.
- Setting goals/objectives – involves identifying a learning goal or objective regarding a topic being considered in class.
- Student discussion/chunking – involves breaking a lesson into chunks for student or group discussion regarding the content being considered.
- Summarizing – involves requiring students to provide a brief summary of content.
- Tracking student progress and scoring scales – involves the use of scoring scales and tracking student progress toward a learning goal.

Figure 7. Random Effects for Specific Instructional Strategies

Instructional Strategy	N	$\overline{ES}$	SE	95% CI		Sig. (2-tailed)	% Gain
				LL	UL		
Advance Organizers	7	.03(.04)	.23(.26)	-.43(-.48)	.49(.56)	.899(.886)	1(2)
Building Vocabulary	41	.44(.51)	.10(.11)	.25(.29)	.64(.73)	.000(.000)	17(20)
Effort and Recognition	11	.31(.37)	.20(.23)	-.09(-.08)	.71(.82)	.130(.107)	12(14)
Feedback	7	.10(.11)	.24(.27)	-.38(-.42)	.57(.64)	.687(.687)	4(4)
Graphic Organizers	65	.29(.34)	.08(.09)	.13(.16)	.44(.51)	.000(.000)	11(13)
Homework	8	.33(.38)	.23(.26)	-.12(-.12)	.78(.88)	.149(.138)	13(15)
Identifying Similarities and Differences	52	.46(.52)	.09(.10)	.28(.33)	.63(.72)	.000(.000)	18(20)
Interactive Games	62	.46(.53)	.08(.09)	.30(.35)	.62(.71)	.000(.000)	18(20)
Nonlinguistic Representations	129	.38(.44)	.06(.06)	.27(.32)	.49(.56)	.000(.000)	15(17)
Note Taking	46	.38(.44)	.09(.10)	.20(.24)	.56(.64)	.000(.000)	15(17)
Practice	5	.32(.37)	.29(.32)	-.25(-.26)	.89(1.01)	.266(.251)	13(14)
Setting Goals/Objectives	16	.57(.66)	.16(.18)	.26(.31)	.89(1.02)	.000(.000)	22(25)
Student Discussion/Chunking	53	.37(.43)	.09(.10)	.20(.23)	.54(.62)	.000(.000)	14(17)

Instructional Strategy	N	$\overline{ES}$	SE	95% CI		Sig. (2-tailed)	% Gain
				LL	UL		
Summarizing	17	.42(.49)	.15(.17)	.13(.15)	.72(.82)	.005(.004)	16(19)
Tracking Student Progress and Scoring Scales	14	.87(1.00)	.17(.20)	.53(.62)	1.21(1.39)	.000(.000)	31(34)

Note: See discussion of Figure 2 for a description of column headings. Corrected findings are presented in parentheses.

Figure 8. Homogeneity Analysis for Instructional Strategies

Q	Sig. (2-tailed)	df
16.324 (16.813)	.294 (.266)	14

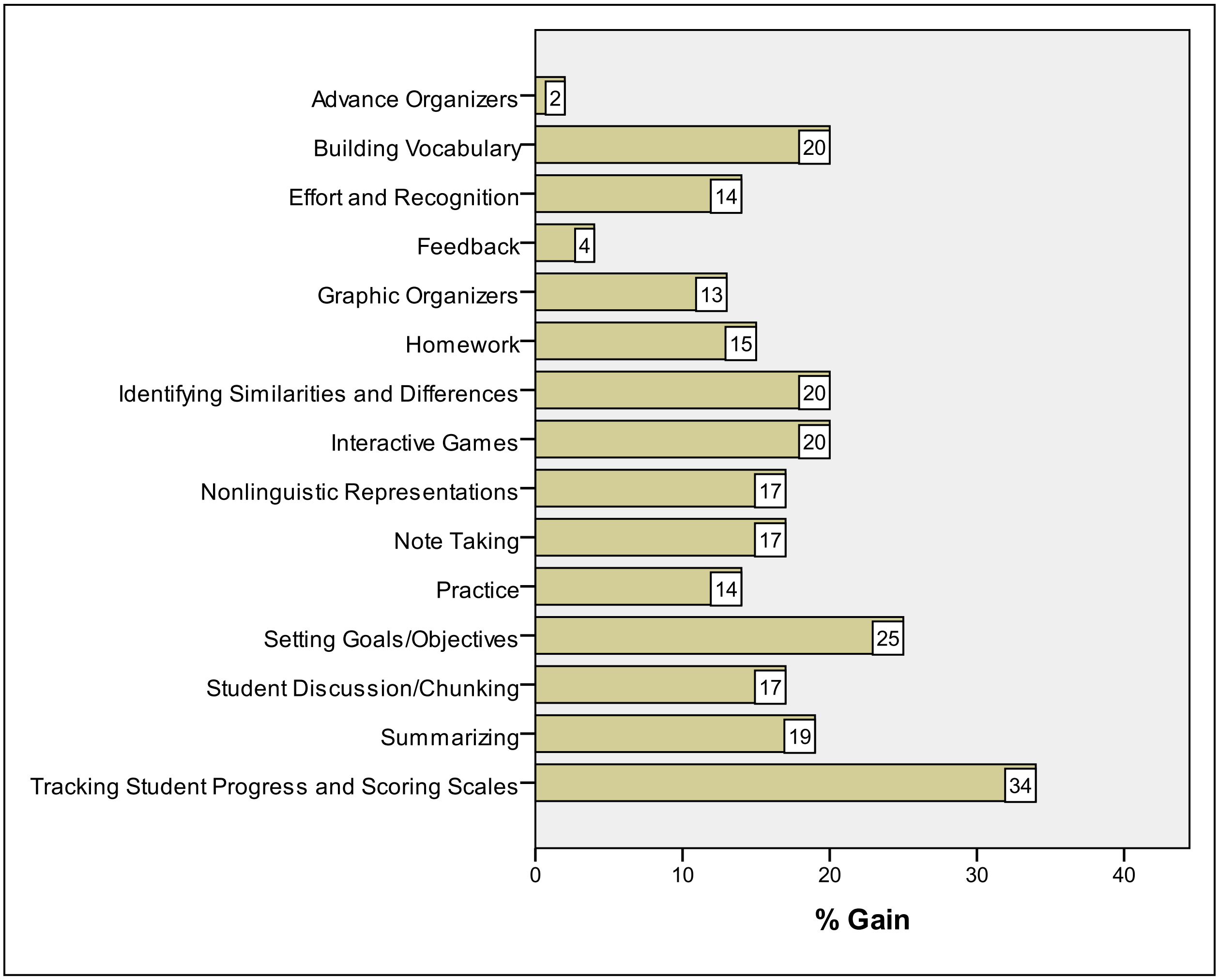
Note: Corrected findings are presented in parentheses.

Figure 7 shows the random-effects estimate for the 15 instructional strategies. Some of the 329 independent studies were included in the meta-analysis for more than one strategy. This occurred when one instructional strategy was a subcomponent of another strategy. For example, the strategy of nonlinguistic representations is also a subcomponent of the strategy for building vocabulary.

The weighted average effect sizes reported in Figure 7 were statistically significant at the .0001 level ($p < .0001$) for seven instructional strategies (building vocabulary, identifying similarities and differences, interactive games, nonlinguistic representations, note taking, student discussion/ chunking, tracking student progress and scoring scales), at the .001 level ($p < .001$) for two instructional strategies (graphic organizers, setting goals/objectives), and at the .01 level ($p < .01$) for one instructional strategy (summarizing). The associated percentile gain was positive for all 15 instructional strategies. As indicated in Figure 8 the homogeneity analysis for instructional strategies was not statistically significant ($p < .05$). Taken at face value this would indicate that the effect sizes all come from the same population.

Figure 11 graphically depicts the percentile gains associated with each instructional strategy.

Figure 11. Percentile Gain for Specific Instructional Strategies (Corrected)



Interpretation

There are a number of ways to interpret an effect size. One interpretation is the amount of overlap between the experimental and control groups. Consider again that an effect size of 1.00 can be interpreted as the average score in the experimental group being one standard deviation higher than the average score in the control group. Consulting a table of the normal curve (i.e., normal distribution) the associated percentile gain for an effect size of 1.00 is 34. This means that the score of the average student in the experimental group (50^{th} percentile) exceeds the scores of 84 percent of the control group. Only 16 percent of the control group would be expected to have scores that exceed the score of the average student in the experimental group.

Figure 10 depicts the percentage of control group students who scored lower than the average student in the experimental group (50^{th} percentile). When corrected for attenuation, the average student in the experimental group (i.e., the group that used an instructional strategy) scored higher than 66% of the students in the control group (i.e., the group that did not use an instructional strategy).

Figure 10. Amount of Overlap between Experimental and Control Groups

	$\overline{ES}$	**Percentage of Control Group Scoring Lower than Experimental Average (50^{th} Percentile)**
Overall	.36 (.42)	64% (66%)

Note: Corrected findings are presented in parentheses.

Another interpretation is to consider the hypothetical change in rank for a class with 100 students. Figure 11 displays this interpretation.

Figure 11. Hypothetical Change in a Student's Class Rank

Class Rank Without Instructional Strategies	Average Student's Class Rank With Instructional Strategies
1	1
5	5
10	10
15	15
20	20
25	25
30	30
35	**34**
40	35
45	40
50	45
55	50
60	55
65	60
70	65
75	70
80	75
85	80
90	85
95	90
100	95

Figure 11 shows the hypothetical change in class rank of the average student in the control group (50^{th} percentile). If that student were the only student to receive instruction using these strategies, his or her class rank would be expected to increase from 50^{th} to 34^{th}.

Summary

This meta-analysis sought to answer the following questions:

1. What effect does the utilization of instructional strategies have on students' achievement regarding the subject matter content taught by their teachers?
2. Does the effect of instructional strategies differ between school levels?
3. Does the effect of instructional strategies differ from strategy to strategy?

The average effect size for all 329 independent studies was statistically significant ($p < .0001$). When corrected for attenuation, the percentile gain associated with the use of the instructional strategies is 16 ($\overline{ES} = .42$). This means that on the average, the strategies used in the independent studies represent a gain of 16 percentile points over what would be expected if teachers did not use the instructional strategies. A reasonable inference is that the overall effect of a 16 percentile point gain is probably not a function of random factors that are specific to the independent studies; rather, the 16 percentile point increase represents a real change in student learning.

Technical Notes

Technical Note 1: Conceptually, analysis of covariance (ANCOVA) can be loosely thought of as using the covariate (i.e., pretest score) to predict students' performance on the posttest and then using the residual score (i.e., predicted score minus observed score) for each student as the dependent measure. To illustrate, consider an independent action research study for a topic within mathematics. Using ANCOVA, students' posttest scores were predicted from the scores received on the pretest. The difference between the predicted posttest scores and the observed posttest scores was then computed for each student that took both pretest and posttest. This difference is referred to as the residual score for each student. It represents the part of each student's posttest score that cannot be predicted from the pretest score for that student. Theoretically, use of residual scores based on pretest predictions is an attempt to equate all students on the dependent measure prior to execution of the intervention—in this case the use of the target instructional strategy (e.g., vocabulary).

Technical Note 2: The meta-analytic findings in this report are typically reported in two ways, 1) findings based on the observed effect sizes from each independent study (see Appendix B), and 2) findings based on a correction for attenuation due to lack of reliability in the dependent measure (i.e., teacher designed assessments of student academic achievement). Hunter and Schmidt detail the rationale and importance of correcting for 11 attenuation artifacts—one of which is random error associated with measurement of the dependent variable (2004, pp. 301-313). They explain:

> . . . error of measurement in the dependent variable reduces the effect size estimate. If the reliability of measurement is low, the reduction can be quite sizable. Failure to correct for the attenuation due to error of measurement yields an erroneous effect size estimate. Furthermore, because the error is systematic, a bare-bones meta-analysis on uncorrected effect sizes will produce an incorrect estimate of the true effect size. The extent of the reduction in the mean effect size is determined by the mean level of reliability across the studies. Variation in reliability across studies causes variation in the observed effect size above and beyond that produced by sampling error. . . . A bare-bones meta-analysis will not correct for either the systematic reduction in the mean effect size or the systematic increase in the variance of effect sizes. Thus, even meta-analysis will produce correct values for the distribution of effect sizes only if there is a correction for the attenuation due to error of measurement. (p. 302)

For ease of discussion consider correcting for attenuation due to unreliability in the dependent measure using the population correlation instead of the population standardized mean difference effect size. The reader should note that the example provided regarding correcting correlations is analogous to correcting a standardized mean difference. To illustrate correcting for attenuation due to unreliability in the dependent measure, assume that the population correlation between the target instructional strategy (e.g., nonlinguistic representations) and student academic achievement is .50. A given study attempts to estimate that correlation but employs a measure of the dependent variable (i.e., a teacher designed assessment of student academic achievement) that has a reliability of .81—considered a typical reliability for a test of general cognitive ability. According to attenuation theory, the correlation would be reduced by the square root of the reliability (i.e., the attenuation factor). In other words, the population correlation is multiplied by the attenuation factor ($\sqrt{.81}$ = .90), thus reducing the correlation by 10 percent. Therefore, the observed correlation will be .45 (.50 x .90) even if there is no attenuation due to the other ten artifacts listed by Hunter and Schmidt (2004, p. 35). When the measure of the dependent variable has a lower reliability, .36 for example, the correlation is reduced by 40 percent ($\sqrt{.36}$ = .60) to .30 (.50 x .60). In order to make a correction for attenuation, the correlation is divided by the attenuation factor (i.e., the square root of the reliability).

For the purposes of this report, an estimate of reliability was used. Osborne (2003) found that the average reliability reported in psychology journals is .83. Lou and colleagues (1996) report a typical reliability of .85 for standardized achievement tests and a reliability of .75 for unstandardized achievement tests. Because the dependent measure in the independent studies involved teacher-designed assessments of student academic achievement, .75 was used as the reliability to correct for attenuation using the following formula:

$$d_c = \frac{d}{a}$$

In the formula, d_c is the corrected effect size, d is the observed effect size, and a is the attenuation factor (the square root of the reliability). Using this formula, each effect size reported in Appendix B was corrected for attenuation to produce the corrected meta-analytic findings considered in this report.

Technical Note 3: Independent variables can be analyzed as fixed effects or as random effects. In the context of ANOVA/ANCOVA, fixed effects are factors that are deliberately arranged by the researcher. In the case of the original analysis of the 329 independent studies, the experimental/ control condition (i.e., the use of a selected instructional strategy) was analyzed as a fixed effect. In contrast, random effects are factors that are not deliberately arranged. Instead, random effects are factors which are randomly sampled from a population of possible samples. Generally speaking, when independent variables are analyzed as random effects, the intent is to generalize results beyond the boundaries of the independent variables employed in the study. For example, if a researcher were interested in the effect that the quality of school leadership has on academic proficiency, the researcher could select a random sample of schools in order to estimate the amount of variance in student academic achievement attributable to differences between types of school leaders. Thus, using the sample, the researcher can make generalizations regarding the influence of school leadership on academic achievement as a whole. Additional research could attempt to replicate the findings by selecting a different random sample of schools for comparison. When fixed effects are employed one typically does not generalize beyond the boundaries of the independent variables in the study. Because the experimental versus control condition in the independent studies was considered a fixed effect, generalizations should be considered with caution as they can be made only with respect to the use of instructional strategies by teachers involved in the independent studies.

Technical Note 4: In Appendix B, the column labeled "ES" contains the computed effect size for each study calculated as Cohen's δ using the following formula:

$$d = \frac{r}{\sqrt{(1 - r^2)(p(1 - p))}}$$

where r is the effect size correlation and p is the proportion of the total population in one of the two groups (i.e., the experimental group). Partial eta squared (η_p^2) as calculated by SPSS was used to determine partial eta (η_p) as an estimate for r by taking its square root. This formula is used to compute the effect size from an effect size correlation (e.g., the point-biserial correlation coefficient) when the experimental and control group populations are not equal (see Lipsey & Wilson, 2001, pp. 62-63). Again, partial eta (η_p) was used as an estimate for r in the formula.

The generic term *effect size* applies to a variety of indices (e.g., *r, R, PV*) that can be used to demonstrate the effect of an independent variable (e.g., use of a selected instructional strategy) on a dependent variable (e.g., student academic achievement). In this report, the effect size statistic utilized is the standardized mean difference effect size. This index, first popularized by Glass (1976) and Cohen (1977), is the difference between experimental and control means divided by an estimate of the population standard deviation.

$$\textit{standardized mean difference effect size} = \frac{\textit{mean of experimental group} - \textit{mean of control group}}{\textit{estimate of population standard deviation}}$$

Consider the following illustration of the use of effect size. Assume that the achievement mean of a group of students in a class that used a target instructional strategy (e.g., graphic organizers) is 90 on a standardized test and the mean of a group of students in a class that did not use the instructional strategy is 80. Assuming the population standard deviation is 10, the effect size would be as follows:

$$ES = \frac{90 - 80}{10} = 1.0$$

This effect size leads to the following interpretation: The mean of the experimental group is 1.0 standard deviation larger than the mean of the control group. One could infer from this that the use of graphic organizers raises achievement test scores by one standard deviation. Therefore, the effect size expresses the differences between means in standardized or "Z score" form, which gives rise to another index frequently used in research regarding education—percentile gain.

Percentile gain is the expected gain (or loss) associated with the effect size expressed in percentile points of the average student in the experimental group compared to the average student in the control group. By way of illustration, consider the same example. An effect size of 1.0 can be interpreted as the average score in the experimental group being about 34 percentile points greater than the average score in the control group. Again, the effect size translates the difference between group means into Z score form. Distribution theory dictates that a Z score of 1.0 is at the 84.13 percentile point of the standard normal distribution. To determine the percentile gain, the effect size is transformed into percentile points above or below the 50^{th} percentile point on the unit normal distribution (e.g., 84% - 50% = 34%).

Technical Note 5: Within the context of meta-analysis, independent studies can be analyzed using a fixed-effect or random-effect model of error to calculate the variability in effect size estimates averaged across the studies. Fixed-effect models calculate error that reflects variation in studies' outcomes due to the sampling of participants (i.e., sampling error) alone. In contrast, random-effect models allow for the possibility that, in addition to sampling error, the effect size varies from study to study due to variations in study methods. Stated differently, random-effect models make an assumption that study-level variance is present as an additional source of random influence. (For a more thorough discussion regarding models used in meta-analysis, see Hunter & Schmidt, 2004; Lipsey & Wilson, 2001; Cooper, 2009.)

Appendix A: Instructions for Action Research

Thank you for agreeing to participate in an action research study regarding the effectiveness and utility of instructional strategies in your classroom. To be involved in a study you must be willing to do a few things. First, you should select a specific unit of instruction, or set of related lessons on a single topic (hereinafter referred to as unit) and design a pretest and posttest for that unit. It is best if the unit is relatively short in nature. For example, if you teach mathematics, you might select a one week unit on linear equations. Second, you must deliver the same unit to two different groups of students (a experimental group and a control group).

At the beginning of the unit, you would administer a pretest on linear equations. Then at the end of the unit you would administer a posttest. This test could be identical to the pretest, or it could be different. The important point is that you have a pretest and a posttest score for each student on the topic of linear equations. The pretest and posttest should be comprehensive in nature. Also, you would administer the same pretest and posttest to both experimental and control groups.

Again, you are teaching the same unit to two different groups of students (experimental and control). Ideally, you would teach the unit to both groups during the same period of time. When teaching the unit of instruction to the experimental group, you would make sure you use your target instructional strategy whenever and in ways you believe it to be applicable. When teaching the unit of instruction to the control group, you would NOT use your target instructional strategy.

If you are an elementary school teacher and do not have two different classes of students then you would teach two different units within the same subject area to the same students. For example, you might select the subject area of writing. First, you might teach a one week unit of instruction on writing essays that focus on logical progression of ideas with good transition sentences. You would begin the unit with a pretest composition that is scored using a rubric specifically designed to measure students' logical progression of ideas and use of good transition sentences. At the end of the unit you would assign another composition, this one used as a posttest. Again, you would score the composition using the same rubric. During this unit of instruction, you would make sure you use your target instructional strategy whenever and in ways you believe it to be applicable. Then, you might teach a one week unit of instruction on writing essays with a clear purpose for a specific audience. As before, you would begin the unit with a pretest composition that is scored using a rubric specifically designed to measure students' presentation of a clear purpose for a specific audience. At the end of the unit you would assign another composition, this one used as a posttest. Again, you would score the composition using the same rubric. During this unit of instruction you would NOT use your target instructional strategy.

Pretest and posttest scores for each student would be recorded on the appropriate form (see below for sample forms), along with general demographic information for each student. If a student does NOT take a test, leave a blank space on the form to indicate a missing test. Please note there is no space for including student names or other means of identifying each student. This has been done intentionally to comply with student privacy requirements. This is an anonymous action research study; do NOT include any student names, id numbers, or other student identifiers on the data sheets you submit to Marzano Research Laboratory. Both pretest and posttest scores should be translated to a percentage format without the percentage sign (i.e. 90% = 90). For example, if your pretest involves 20 points and a particular student receives a score of 15, then translate the 15 into a percentage of 75 (i.e. 15/20 = .75 x 100 = 75) and record that as the pretest score for the student. If your posttest involves 80 points and that same student receives a score of 75, then translate the 75 into a percentage of 94 (75/80 = .94 x 100 = 94) and record that as the student's posttest score. The same procedure would be employed if you used a rubric. For example, if a student received a 2 on a 4 point rubric on the pretest, this score would be translated to a percentage of 50 (2/4 = .50 x 100 = 50) and this would be recorded as the student's pretest score. The same translation would be done on the student's rubric score for the posttest. Again, leaving the percentage sign off the score recorded on the forms.

It is imperative that you keep track of each student's pretest scores and posttest scores and make sure they match when your data sheet is filled out. If posttest scores are not aligned with the pretest scores for particular students then the data cannot be used.

When you have completed the study please fill out the required forms and return them to your team leader for submission to Marzano Research Laboratory. Three separate forms are required. The first is a brief survey form which asks you to provide general information about your action research study, your target instructional strategy, and your experience as a teacher. The remaining forms ask you to provide anonymous demographic information about your students along with their pretest and posttest scores. One form is for students in the experimental group, i.e. the students in the group that used the target instructional strategy. The other form is for students in the control group, i.e. the students that did NOT use the target instructional strategy. Please use the ethnicity codes listed at the bottom of each form when filling out the demographic information for your students.

Thank you again for considering involvement in an action research project.

Name ___________________________________

School (optional) ___________________________

District (optional) ___________________________

Grade level(s) taught__________________________

Target Instructional Strategy________________________

Topic (and general subject area) addressed during the unit where the target instructional strategy was used (experimental group) ________________________________

Unit length (# of days) ___________

Topic (and general subject area) addressed during the unit where the target instructional strategy was NOT used (control group) ________________________________

Unit length (# of days) ___________

Were both classes comprised of different students? (Y/N) _______

General description of what you did - Target Instructional Strategy Class (Experimental Group):

General description of what you did - Non-Target Instructional Strategy Class (Control Group):

Experimental Group Scores – Target Instructional Strategy Used

Student	Grade	Gender	Ethnicity	Free/Reduced Lunch (Y/N)	English Language Learner (Y/N)	Special Education (Y/N)	Pretest Score	Posttest Score
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								

Ethnicity Code: A – Asian, AA – African American, C – White/Caucasian, H – Hispanic, N – Native American, O – Other

Control Group Scores **–** Target Instructional Strategy **NOT** Used

Student	Grade	Gender	Ethnicity	Free/Reduced Lunch (Y/N)	English Language Learner (Y/N)	Special Education (Y/N)	Pretest Score	Posttest Score
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								

Ethnicity Code: A – Asian, AA – African American, C – White/Caucasian, H – Hispanic, N – Native American, O - Other

Teacher Survey

How long have you been teaching? __________

How long have you used your target instructional strategy in your classroom? ____________

How confident are you in your ability to use your target instructional strategy in your classroom?

Not at all				Completely
1	2	3	4	5

Appendix B: Independent studies

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
1	9-12	Collaborative Research	24	21	1.08	.001	36
2	9-12	Visual vs verbal instruction	18	15	2.22	.000	49
3	9-12	Graphic organizers	25	27	.68	.022	25
4	9-12	Systematic homework feedback	13	13	-.45	.290	-17
5	9-12	Recalling/activating prior knowledge	21	19	.16	.639	6
6	9-12	Note taking	28	30	2.39	.000	49
7	9-12	Using formatted note sheet	31	30	-.25	.337	-10
8	9-12	Unidentified	16	15	-.20	.594	-8
9	9-12	Nonlinguistic representations	20	16	.77	.034	28
10	9-12	Nonlinguistic representations	19	20	.62	.069	23
11	9-12	Nonlinguistic representations	13	13	-.65	.130	-24
12	9-12	Nonlinguistic representations	13	12	.06	.898	2
13	9-12	Homework	26	20	.03	.920	1
14	9-12	Nonlinguistic	23	28	.11	.696	4
15	9-12	Reinforcing effort	17	23	-.29	.383	-11
16	9-12	Summarizing	25	28	.31	.284	12
17	9-12	Nonlinguistic	17	24	1.97	.000	48
18	9-12	Summarizing	27	22	.46	.124	18
19	9-12	Reinforcing effort	25	27	-.51	.079	-20
20	9-12	Reinforcing effort	15	18	.13	.735	5
21	9-12	Nonlinguistic	17	20	1.06	.004	36
22	9-12	Nonlinguistic	20	20	.26	.423	10
23	9-12	Summarizing	23	25	.00	.887	0
24	9-12	Summarizing	24	29	.21	.452	8
25	9-12	Summarizing	26	29	.30	.286	12
26	9-12	Nonlinguistic	21	13	-.16	.668	-6

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
27	9-12	Summarizing	25	16	-.06	.841	-2
28	9-12	Summarizing	15	27	.95	.005	33
29	9-12	Summarizing	22	20	.30	.356	12
30	9-12	Cues, Questions and Advance Organizers	27	26	.06	.821	2
31	9-12	Homework	15	14	.34	.390	13
32	9-12	Basics for 2x2, 3x3 matrices	16	14	.62	.120	23
33	9-12	Nonlinguistic	22	20	.28	.390	11
34	9-12	Nonlinguistic	23	20	.24	.460	9
35	9-12	Nonlinguistic	22	19	.68	.040	25
36	9-12	Comparisons	12	30	-.40	.220	-16
37	9-12	Nonlinguistic	20	14	1.30	.000	40
38	9-12	Computer based Instruction	6	13	.01	.980	0
39	9-12	Homework	26	18	-.32	.320	-13
40	9-12	Comparisons	18	24	-.16	.630	-6
41	9-12	Reinforcing Effort	26	28	-.66	.020	-25
42	9-12	Nonlinguistic	19	19	-.20	.570	-8
43	9-12	Comparisons	19	19	-.61	.080	-23
44	9-12	Nonlinguistic	11	16	.89	.040	31
45	9-12	Nonlinguistic	4	4	1.77	.100	46
46	9-12	Nonlinguistic	14	9	.33	.480	13
47	9-12	Nonlinguistic	12	6	.30	.570	12
48	9-12	Homework	28	26	1.53	.000	44
49	9-12	Reinforcing Effort	8	8	3.11	.000	50
50	9-12	Nonlinguistic	11	13	.31	.480	12
51	9-12	Nonlinguistic	7	7	.88	.170	31
52	9-12	Cooperative Learning	6	6	4.27	.000	50
53	9-12	Nonlinguistic	4	7	1.29	.110	40

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
54	9-12	Reinforcing Effort	3	3	2.53	.120	49
55	9-12	Nonlinguistic	7	6	-.36	.580	-14
56	9-12	Reinforcing Effort	3	6	.97	.280	33
57	9-12	Unidentified	16	19	.17	.640	7
58	9-12	Building vocabulary	10	9	.00	.960	0
59	9	Vocabulary Notebook	14	13	.39	.350	15
60	9	Vocabulary Notebook	4	11	.00	.966	0
61	9-12	Vocabulary Notebook	15	15	1.54	.010	44
62	9-12	Vocabulary	18	17	.31	.380	12
63	10-12	Six Steps of Vocabulary	16	11	-.41	.180	-16
64	9	Vocabulary Notebook	4	12	1.68	.010	45
65	10-12	Vocabulary Notebook	11	11	.79	.070	29
66	9-12	Generating Hypotheses	27	27	-.45	.200	-17
67	9	Similarities and Differences	23	22	.36	.249	14
68	10	Vocabulary Notebook	20	22	.86	.010	31
69	11	Vocabulary Notebook	27	28	.00	.899	0
70	9	Vocabulary Notebook	25	17	.16	.634	6
71	9-12	Graphic Organizers	10	12	.13	.790	5
72	9-12	Graphic Organizer	7	7	-.26	.681	-10
73		Vocabulary Notebook	8	8	2.27	.001	49
74	11	Vocabulary Notebook	18	25	-.63	.054	-24
75	10-12	Graphic Organizers	22	27	.11	.711	4
76	9	Graphic Organizers	25	18	-.11	.752	-4
77		Graphic Organizers	23	24	.62	.045	23
78	9	Building Academic Vocabulary	15	19	.71	.056	26
79	9-12	Nonlinguistic Depictions	23	25	.20	.510	8
80	9-12	Summarizing/Note Taking	18	20	.90	.011	32

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
81	9	Graphic Organizers	17	20	.14	.680	6
82	9	Graphic Organizers	17	19	.16	.666	6
83	10	Note Taking	25	24	-.06	.860	-2
84	10	Note Taking	24	25	-.31	.292	-12
85	9	Building Vocabulary/Images Strategy	19	26	1.23	.000	39
86	9	Homework and Practice	16	22	.20	.564	8
87	9	Note Taking	19	21	.09	.810	4
88	10-12	Graphic Organizer/Use of pre test	27	27	.49	.090	19
89	9-11	Vocabulary Notebook	5	8	.25	.700	10
90	11-12	Vocab Notebook	21	25	-.13	.670	-5
91	11	Note Taking	22	18	.37	.270	14
92		Review and Practice	20	20	.32	.332	13
93	9	Interactive Games	9	10	1.54	.010	44
94	.	Practice and Feedback	17	19	.31	.380	12
95	9	Note Taking	31	16	-.41	.180	-16
96	10	Quia Word Games	29	21	1.23	.000	39
97	10	Note Taking	10	16	.79	.070	29
98	11	Note Taking	20	17	-.45	.200	-17
99	12	Nonlinguistic Representations	15	14	-.57	.160	-22
100	11	Graphic Organizers	28	21	-.80	.010	-29
101	11	Building vocabulary	29	32	.42	.118	16
102	10	Blue Clickers	21	24	.24	.451	9
103	9	Graphic Organizers	22	19	-.06	.824	-2
104	12	Vocab G. Org. and Games	26	25	.47	.110	18
105		Vocab Review Games	11	15	-.26	.551	-10
106	9	Vocabulary Review Games	26	25	.64	.031	24
107	10-12	Graphic Organizers	17	19	.53	.141	20

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
108	9-12	Graphic Organizers	14	21	.76	.039	28
109	10-12	Table Talk	24	22	.73	.022	27
110	10	Pyramid Game	26	26	.44	.131	17
111	9-11	Graphical Organizer (roundhouse diagram)	10	17	.00	.997	0
112	9-11	CPS Instruction	11	8	.19	.703	8
113	11	Voc. Step 3 construct a graphic	5	11	1.06	.079	36
114	9-12	Review Games	9	22	.56	.151	21
115	10-12	Practice - Feedback	10	13	.00	.977	0
116	10	Note Taking		25	.66	.001	25
117	.	Note Taking	12	8	.64	.200	24
118	.	Vocab: Poster/G. Org.	13	5	1.02	.070	35
119	10	Vocab	7	12	-.94	.080	-33
120	10	Vocab. 6 Strategies	17	16	.31	.400	12
121	9-12	Quia Word Games	16	22	-.14	.677	-6
122	.	Vocab - Games	23	14	.30	.390	12
123	10	Note Taking	4	4	-2.40	.040	-49
124	9-12	Graphic Organizer	15	44	.29	.280	11
125	9-12	Nonlinguistic	25	18	.93	.000	32
126	9-12	2 Column Notes with Picturing	24	24	-.19	.530	-8
127	9-12	Nonlinguistic	26	13	.84	.020	30
128	9-12	Nonlinguistic	19	21	-.39	.240	-15
129	9-12	Concept Pattern Organizer	22	20	.06	.950	2
130	9-12	Comparison Matrix and Recategorization	39	43	-.13	.560	-5
131	9-12	Compare and Contrast	26	22	.72	.020	26
132	9-12	Nonlinguistic	21	25	-.39	.210	-15
133	9-12	Student Notes: Combination Technique	23	27	.06	.950	2
134	9-12	Notes and Summarizing	24	17	1.59	.000	44

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
135	9-12	Nonlinguistic	57	57	-.09	.240	-4
136	10	Nonlinguistic	23	23	.25	.430	10
137	9-12	Engaging in Kinesthetic Activity	3	4	1.60	.190	45
138	11	Nonlinguistic	15	13	-.07	.920	-3
139	9-12	Historical Inquiry	23	14	-.93	.010	-32
140	9-12	Reinforcing Effort	19	19	-.93	.010	-32
141	10	Nonlinguistic	45	89	.24	.180	9
142	9-12	Nonlinguistic	20	15	.51	.160	20
143	9-12	Cornell Note-taking System	40	41	.53	.020	20
144	9-12	Nonlinguistic/$20,000 pyramid	36	32	.06	.920	2
145	9-12	Combination Notes	13	25	.33	.330	13
146	9-12	Nonlinguistic	15	18	.06	.910	2
147	9-12	Nonlinguistic	22	27	.25	.410	10
148	9-12	Nonlinguistic	36	61	-.06	.870	-2
149	8	Nonlinguistic Representations	9	18	1.46	.002	43
150		Nonlinguistic Representation	24	18	-.17	.598	-7
151		Note taking	22	66	-.51	.022	-20
152	6	"I" Chart Note Taking	15	63	-.06	.765	-2
153	8	Note Taking	24	103	.64	.001	24
154		Graphic Organizer	40	41	.37	.106	14
155	6	Graphic Organizer	20	65	-.29	.195	-11
156	7	Notes w/drawings	21	69	.36	.091	14
157	8	Note Taking	20	83	.16	.427	6
158		Nonlinguistic Representations	24	24	-.21	.484	-8
159	8	Notetaking/Linguistc/Nonlinguistic	22	23	1.11	.001	37
160	7	Nonlinguistic Representations	26	64	.24	.268	9
161		Graphic Organizer	22	60	-.27	.239	-11

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
162	8	Illustrated Terms	23	66	-.09	.693	-4
163		Identifying Similarities/Differences	24	20	1.75	.000	46
164		Effort and Recognition	16	15	.60	.125	23
165		Nonlinguistic Representations	20	29	-.32	.279	-13
166	7	Graphic Organizer	41	38	-.21	.350	-8
167	6	Nonlinguistic Representation	24	70	.16	.453	6
168	8	Graphic Organizer	20	44	-.16	.557	-6
169		Nonlinguistic	21	20	-.47	.156	-18
170	6	Graphic Organizer	23	59	.06	.788	2
171		Nonlinguistic note taking	16	71	.82	.000	29
172		Graphic Organizer	19	59	.00	.933	0
173		Effort	25	23	2.28	.000	49
174		Effort and Achievement	60	60	.00	.959	0
175		Note Taking	22	66	.50	.024	19
176		Note Taking	25	72	.33	.113	13
177	8	Notetaking Nonlinguistic/Linguistic	25	24	.35	.251	14
178	8	Note Taking Strategies	24	64	.27	.215	11
179		Kinesthetic Activities with notes (Foldable)	25	42	.32	.204	13
180		Graphic Organizer	24	60	.16	.473	6
181		Double column notes (in pictures/in words)	16	46	1.14	.000	37
182		Graphic Organizers	24	81	.00	.826	0
183	8	Note Taking	18	66	.97	.000	33
184	6	Nonlinguistic Representations	22	43	.09	.748	4
185	7	Setting objectives and providing feedback	21	45	-.17	.519	-7
186		Note Taking	86	22	.32	.103	13
187		Note Taking	22	63	.28	.212	11
188		highlighting with a yellow highlighter important information	23	70	.46	.030	18

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
189	6	Nonlinguistic Representations	20	71	.06	.719	2
190		Non-linguistic representation	22	63	.69	.002	25
191		Nonlinguistic Representations	21	40	.09	.713	4
192		Graphic Organizers	46	42	.00	.937	0
193	6	Nonlinguistic Representation	20	40	.59	.029	22
194		Nonlinguistic Representation	26	23	-.68	.025	-25
195	7	Double column notes with visual and summary statement	23	63	.35	.115	14
196	12	Note Taking	21	21	2.19	.000	49
197	11	Teacher directed PowerPoint Notes	16	29	.19	.537	8
198	11	Teacher directed PowerPoint Notes	25	31	-.26	.353	-10
199	11-12	Nonlinguistic Representations	8	10	.75	.165	27
200	9	Nonlinguistic Representations	13	12	.93	.039	32
201	10-12	Teacher prepared questions to accompany Power Point	19	19	-.24	.489	-9
202	10-12	Partial teacher prepared notes	14	12	.58	.180	22
203	10	Note Taking	21	14	.17	.631	7
204	10	Nonlinguistic Representations	34	48	.09	.677	4
205	9-11	Rule-based summarizing	12	14	-.54	.212	-21
206	11	Nonlinguistic Representations	17	11	1.03	.018	35
207	8	Learning Goal Tracking Folders	15	12	1.50	.001	43
208	7	Homework	6	8	1.32	.053	41
209	10	Chunking	20	20	-.28	.402	-11
210	10	Nonlinguistic representations	16	15	.41	.291	16
211	12	Nonlinguistic representations	13	13	.09	.833	4
212	5	Feedback and tracking	11	15	-.39	.361	-15
213	6	Think-pair-share	13	15	.65	.118	24
214	4	Teacher prepared notes	17	21	.24	.481	9
215	3	Practice	11	13	.81	.079	29

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
216	9-12	Discussion of stated learning goal	10	30	.56	.140	21
217	10	Visualizing & Summarizing	13	8	1.81	.002	46
218	7	Homework	20	21	.28	.399	11
219	6	Setting and discussing objectives	22	18	.02	.961	1
220	7	Setting objectives	16	12	-.26	.527	-10
221	7	Nonlinguistic representations	13	15	.17	.670	7
222	1	Tracking	17	17	.84	.025	30
223	4	Scoring Guide	20	24	2.44	.000	49
224	3	Scoring Guide and Feedback	20	20	3.66	.000	50
225	5	Tracking and Feedback	20	20	.63	.059	24
226	5	Nonlinguistic representations	16	16	-.22	.588	-9
227	2	Basal series matrix	15	20	.72	.051	26
228	1	Learning goal defined and repeated	20	20	1.98	.000	48
229	5	Nonlinguistic representations; Pictionary	18	18	.35	.328	14
230	K	Nonlinguistic representations	17	17	-.51	.170	-20
231	2	Verbalizing learner goals	19	15	-.93	.024	-32
232	3	Verbalizing learning goal	17	17	.32	.400	13
233	K	Verbalizing learning goal, charted progress	13	14	.75	.092	27
234	K	Repeated learning goal, Nonlinguistic representations	30	14	2.18	.000	49
235	8	Tracking learning goal progress	25	29	1.00	.001	34
236	4	Tracking learning goal progress	23	24	.07	.811	3
237	5	Tracking learning goal progress	7	7	1.68	.018	45
238	5	Comparison matrix	18	20	2.07	.000	48
239	6	Tracking learning goal progress	28	28	.07	.800	3
240	4	Teacher notes (Cornell)	27	27	-.58	.044	-22
241	K	Learning goals and feedback	18	18	.66	.067	25
242	5	Teacher notes (Cornell), summarizing	23	23	1.25	.000	39

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
243	3	Nonlinguistic representations	19	25	1.50	.000	43
244	2	Nonlinguistic representations	21	19	-.27	.421	-11
245	1	Tracking	20	22	1.20	.001	38
246	K	Nonlinguistic representations	20	17	.44	.210	17
247	3	Chunking	24	24	.06	.847	2
248	1	Chunking	18	21	.19	.567	8
249	4	Nonlinguistic representations, clozentropy, student questioning	23	21	.56	.082	21
250	2	Graphic organizers	21	15	.66	.070	25
251	7	Tracking	17	17	-.32	.376	-13
252	5	Tracking	25	23	.43	.160	17
253	3	Building vocabulary	16	16	1.53	.000	44
254	1	Building vocabulary	19	20	1.71	.000	46
255		Building vocabulary	16	17	2.38	.000	49
256	3	Building vocabulary	20	22	-.14	.656	-6
257	3	Building vocabulary	16	17	-.47	.207	-18
258	2	Building vocabulary	18	19	-.17	.624	-7
259	K	Building vocabulary	16	8	-.56	.238	-21
260	4	Building vocabulary	16	15	-.41	.286	-16
261	1	Building vocabulary	23	23	.98	.002	34
262	5	Building vocabulary	12	15	1.02	.021	35
263	4	Building vocabulary	20	22	2.19	.000	49
264	4	Building vocabulary	22	23	1.68	.000	45
265	3	Building vocabulary	25	22	-.25	.412	-10
266	K	Building vocabulary	19	19	.48	.165	18
267	5	Building vocabulary	26	27	.63	.029	24
268	2	Building vocabulary	19	20	.47	.166	18
269	5	Building vocabulary (1st 4 steps)	19	19	-.12	.722	-5

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
270	3	Building vocabulary	20	18	-.15	.669	-6
271	3	Building vocabulary	20	19	.16	.642	6
272	2	Building vocabulary	10	10	.16	.749	6
273	1	Building vocabulary	16	20	1.81	.000	46
274	4	Building vocabulary	23	23	-.06	.847	-2
275	2	Building vocabulary	20	17	.40	.254	16
276	3	Building vocabulary	18	17	.81	.028	29
277	7	Building vocabulary (step 6 only)	25	22	.64	.039	24
278	6	Building vocabulary (omitted step 5)	30	30	1.52	.000	44
279	6-8	Building vocabulary (steps 2 & 6 only)	12	9	2.87	.000	50
280	8	Building vocabulary (omitted step 6)	29	17	-.24	.450	-9
281	7	Building vocabulary (omitted step 6)	38	43	-.53	.022	-20
282	7	Building vocabulary (steps 3 & 5 only)	31	22	-.30	.298	-12
283		Building vocabulary	16	18	.40	.273	16
284		Building vocabulary	15	14	.58	.150	22
285		Building vocabulary	24	30	.42	.139	16
286		Building vocabulary	18	25	.13	.695	5
287		Building vocabulary	35	52	.13	.566	5
288	9-12	Building vocabulary	1	5	2.02	.283	48
289	10-12	Building vocabulary	11	32	.03	.940	1
290	11-12	Building vocabulary	13	9	1.27	.014	40
291	9	Illustrations	19	25	-.04	.896	-2
292	9-12	Graphic Organizers	15	15	1.88	.000	47
293	9-10	Vocabulary Cards	20	16	.18	.612	7
294	9	Graphic Organizers	19	15	.49	.185	19
295	11	Graphic Organizers	21	22	.20	.541	8
296	10-12	Graphic Organizers	14	15	.08	.832	3

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
297	10-12	Graphic Organizers	7	10	-.42	.454	-16
298	12	Vocabulary Illustration	24	22	.26	.406	10
299	9-10	Graphic Organizers	8	12	1.15	.032	37
300	10-11	Graphic Organizers	22	18	.11	.744	4
301	9-12	Vocabulary	22	25	.33	.285	13
302	10	Graphic Organizers	11	12	-.24	.599	-9
303	10-12	Graphic Organizers	25	25	.31	.299	12
304	9E	Vocabulary	16	17	.16	.662	6
305	9-12	Graphic Organizers	13	17	.01	.970	0
306	9	Building Background Knowledge	6	6	1.34	.075	41
307	9A	Vocabulary	29	27	-.19	.490	-8
308	10	Graphic Organizers	23	23	.12	.700	5
309	10-11	Vocabulary Notebook	27	22	.87	.005	31
310	11	Graphic Organizer	33	29	.08	.768	3
311	9-11	Vocabulary Review Game	22	17	.46	.176	18
312	10-12	Vocabulary Review Game	12	18	-.19	.632	-8
313	10	Note Taking	20	24	.61	.059	23
314	11-12	Vocabulary Review Game	18	17	.03	.927	1
315	9-11	Graphic Organizer	7	7	-.25	.690	-10
316	10-12	Graphic Organizer	31	28	-.11	.679	-4
317	10	Vocabulary Illustrations	28	29	.40	.150	16
318	10	Graphic Organizer	13	12	1.18	.012	38
319	11-12	Graphic Organizer	13	11	.51	.259	20
320	9,12	Vocabulary Review Game	9	15	.26	.568	10
321	10	Vocabulary Notebook	30	20	.09	.755	4
322	11	Vocabulary Review Game	22	18	.21	.525	8
323	9-12	Note Taking	8	14	-.09	.850	-4

Teacher	Grade	Target Strategy	Ctrl N	Exp N	ES	Sig. (2-tailed)	% Gain
324	9-12	Note Taking	8	9	.97	.091	33
325	10-12	Graphic Organizer	8	16	1.19	.018	38
326	11	Restate/Peer Share	10	15	.32	.473	13
327	9-12	Graphic Organizer	22	25	.13	.665	5
328	10-11	Concept Mapping	16	21	-.12	.723	-5
329	5	Vocabulary Review Games	14	21	1.20	.002	38

Source: *Meta-Analysis Database* (Marzano Research Laboratory, 2009).

References

Cohen, J. (1977). *Statistical power for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Glass, G. V. (1976). Primary, secondary, and meta-analyses of research. *Educational Researcher, 5,* 3-8.

Harlow, L.L., Muliak, A. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press, Inc.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: SAGE Publications.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66* (4), 423-458.

Marzano, R. J. (2004). *Building background knowledge for academic achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano Research Laboratory (2009). *Meta-analysis database.* Retrieved August 24, 2009 from http://www.marzanoresearch.com/research/meta_analysis_database.aspx.

Murphy, K. R., & Myors, B. (2004). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2nd ed.). Mahwah, NJ: Erlbaum.

Osborne, J. W. (2003). Effect sizes and disattenuation of correlation and regression coefficients: Lessons from educational psychology. *Practical Assessment, Research and Evaluation, 8* (11). Retrieved March 20, 2009, from http://PAREonline.net/getvn.asp?v=8&n=11.