# Reading Methods: A Higher-Order Meta-analysis

Maureen Schwartz, Ivy Sherman, and David Weiss

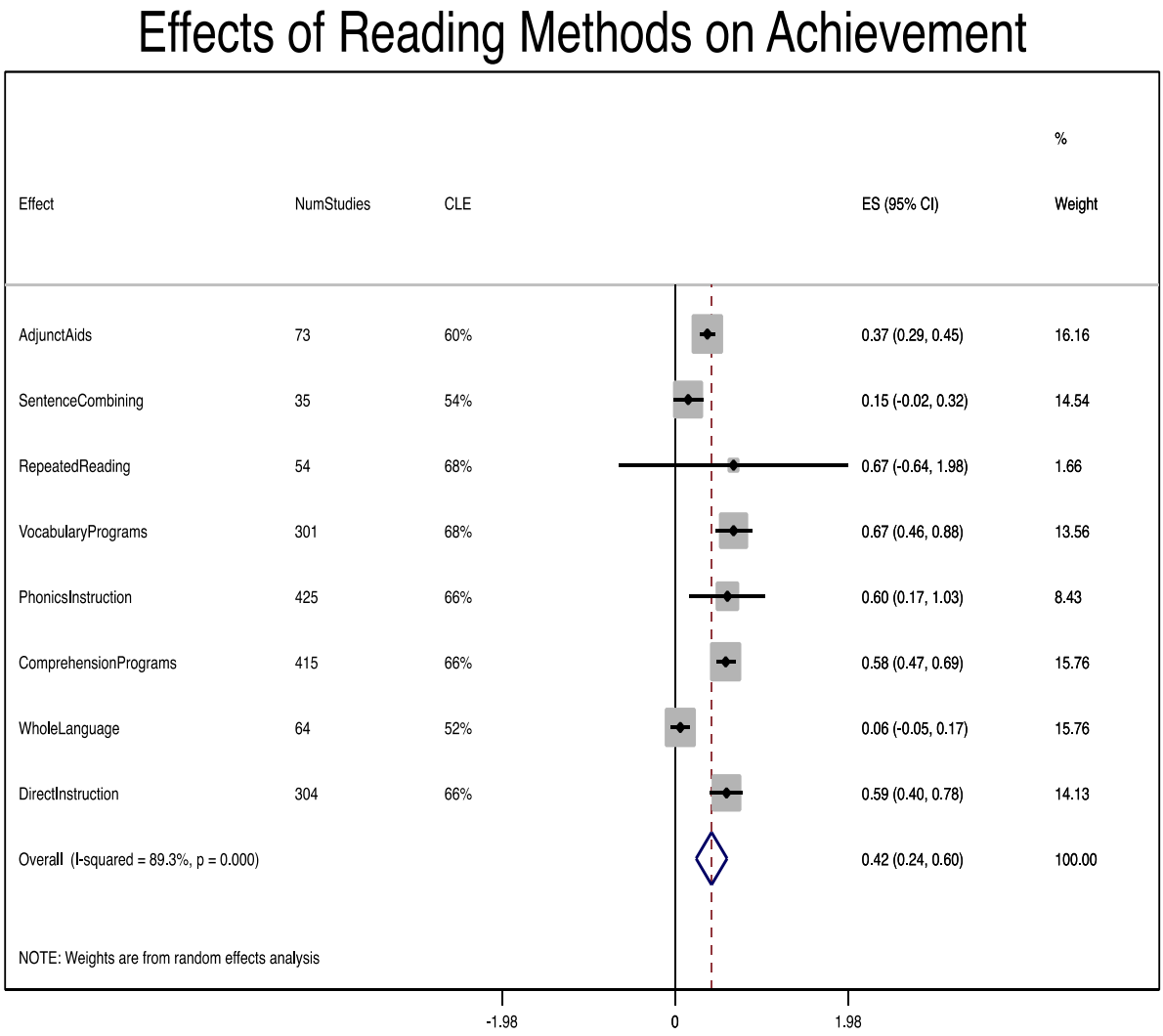
Long Island University/C.W. Post

```
. edit

. save c:\mydata.dta, replace
file c:\mydata.dta saved

. use c:\mydata.dta, clear

. cap set autotabgraphs on

. set more off

. * STEP 2

. * Load data set from

. use c:\mydata.dta

. edit

. * STEP 3

. * Install or update metan, metafunnel, and confunnel ado files for meta-anaysis

. net from http://www.stata-journal.com/software/sj8-1
----------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------
http://www.stata-journal.com/software/sj8-1/
Stata Journal volume 8, issue 1
----------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------

DIRECTORIES you could -net cd- to:
   ..              Other Stata Journals

PACKAGES you could -net describe-:
   gr0031          Spineplots for two-way categorical data
   sbe24_2          Update:  metan: fixed- and random-effects
                  meta-analysis
   sg164_1          Update:  Specification tests for linear panel-data
                  models
   st0126_1         Update:  QIC criterion for model selection in GEE
                  analyses
   st0138          Deterministic and probabilistic sensitivity analysis
                  of epidemiological results
   st0139          A new framework for managing and analysing multiply
```

```
                imputed data in Stata
   st0140       Fuzzy set creation, testing, and reduction
   sxd1_3        Update:  Random allocation of treatments balanced
                in blocks
------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------

. net install sbe24_2
checking sbe24_2 consistency and verifying not already installed...
all files already exist and are up to date.

. net from http://www.stata-journal.com/software/sj4-2
------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------
http://www.stata-journal.com/software/sj4-2/
Stata Journal volume 4, issue 2
------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------

DIRECTORIES you could -net cd- to:
   ..           Other Stata Journals

PACKAGES you could -net describe-:
   gr0002_2      Lean mainstream schemes for Stata 8 graphics
   pr0012       Submenu and dialogs for meta-analysis commands
   sg150_1       Hardy-Weinberg equilibrium test in case-control
                studies
   st0004_2      Update to residual diagnostics for cross-section
                time-series regression models
   st0026_1      Estimation of average treatment effects based on
                propensity scores
   st0030_1      Instrumental variables and GMM:  Estimation and
                testing
   st0058_1      Kaplan-Meier plots with stsatrisk
   st0059       Cumulative incidence estimation in presence of
                competing risks
   st0060       Production function estimation in Stata using inputs
                to control for unobservables
   st0061       Funnel plots in meta-analysis
   st0062       Sample size and power calculations using the noncentral
                t distribution
   st0063       Computing interaction effets and standard errors in
                logit and probit models
   st0064       Confidence intervals for kernel density estimation
------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------
```

```
. net install st0061
checking st0061 consistency and verifying not already installed...
all files already exist and are up to date.

. net from http://www.stata-journal.com/software/sj9-2
----------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------
http://www.stata-journal.com/software/sj9-2/
Stata Journal volume 9, issue 2
----------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------

DIRECTORIES you could -net cd- to:
   ..              Other Stata Journals

PACKAGES you could -net describe-:
   gr0033_1         Update: Contour enhanced funnel plots for
                  meta-analysis
   sbe19_6         Updated tests for bias in meta-analysis
   sbe24_3         Update: metan: fixed- and random-effects
                  meta-analysis
   st0096_2        Update: GLS for trend estimation of summarized
                  dose-response data
   st0139_1        Update: Multiple imputation of missing values
   st0143_2        Update: Fit a linear model with two
                  high-dimensional fixed effects
   st0152_1        Update: Blinder-Oaxaca decomposition for linear
                  and nonlinear models
   st0163          metandi: Meta-analysis of diagnostic accuracy
                  using hierarchical logistic regression
   st0164          Two techniques for investigating interactions
                  between treatment and continuous covariates in
                  clinical trials
   st0165          Further development of flexible parametric models
                  for survival analysis
   st0167          The Skillings-Mack Test (Friedman test when there
                  are missing data)
   st0168          Quasi-Bayes smoothing of categorical frequencies
----------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------

. net install gr0033_1
checking gr0033_1 consistency and verifying not already installed...
all files already exist and are up to date.
```

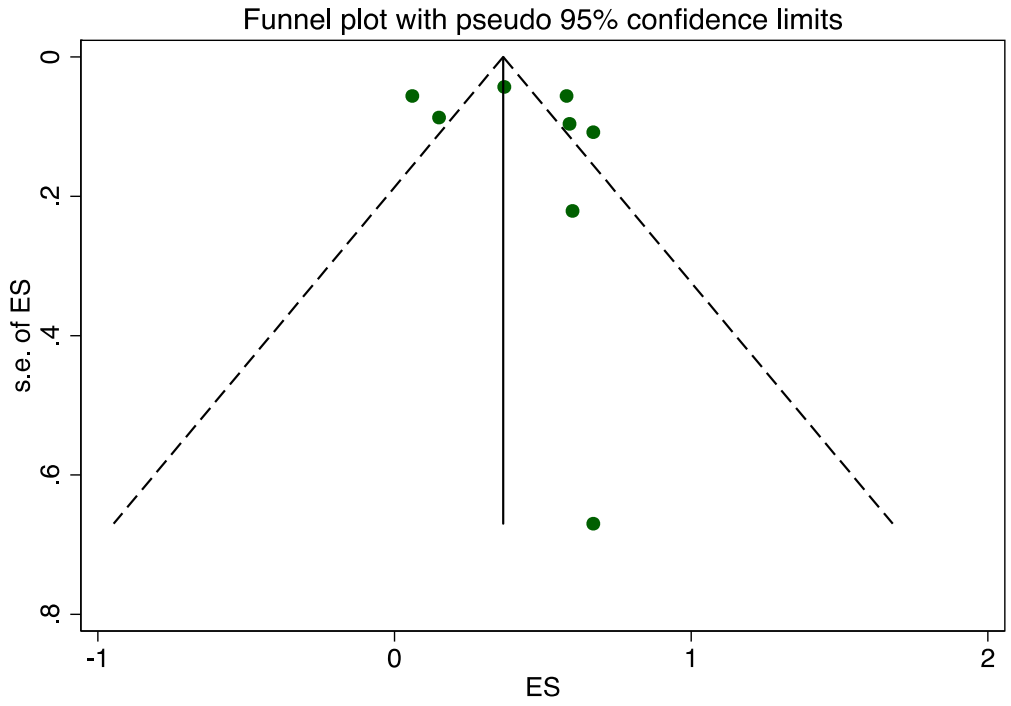
. metan ES SE, random lcols(Effect NumStudies CLE) astext(70) title(Effects of Reading Methods on Achievement) scheme(s1color) note("Based on data from Tables 8.1, p. 130, and 10 > .1, p. 201, VISIBLE LEARNING, J. C. Hattie, 2009")




Based on data from Tables 8.1, p. 130, and 10.1, p. 201, VISIBLE LEARNING, J. C. Hattie, 2009

| Study | ES | [95% Conf. | Interval] | % Weight |
|---|---|---|---|---|
| AdjunctAids | 0.370 | 0.286 | 0.454 | 16.16 |
| SentenceCombining | 0.150 | -0.021 | 0.321 | 14.54 |
| RepeatedReading | 0.670 | -0.643 | 1.983 | 1.66 |
| VocabularyPrograms | 0.670 | 0.458 | 0.882 | 13.56 |
| PhonicsInstruction | 0.600 | 0.167 | 1.033 | 8.43 |
| ComprehensionProgram | 0.580 | 0.470 | 0.690 | 15.76 |
| WholeLanguage | 0.060 | -0.050 | 0.170 | 15.76 |

```
DirectInstruction    |  0.590      0.402     0.778       14.13
---------------------+-------------------------------------------------
D+L pooled ES        |  0.418      0.240     0.597      100.00
---------------------+-------------------------------------------------

  Heterogeneity chi-squared =  65.33 (d.f. = 7) p = 0.000
  I-squared (variation in ES attributable to heterogeneity) =  89.3%
  Estimate of between-study variance Tau-squared =  0.0495

  Test of ES=0 : z=   4.59 p = 0.000

.
. * STEP 5

. * Produce funnel plot.

. metafunnel ES SE, scheme(s1color) name(funnel, replace)

note: default data input format (theta, se_theta) assumed
```




```
. graph save funnel "/Users/teachershortz27/Desktop/funnel plot.gph"
(file /Users/teachershortz27/Desktop/funnel plot.gph saved)
```

Abstract

Following a meta-analysis study conducted by Hattie (2009), the authors carried out a follow-up higher-order meta-analysis of the effect of reading methods on the effectiveness of reading instruction in terms of student achievement. The authors restructured the constructs utilized by Hattie to develop a single effect called reading methods. The researchers analyzed 46 meta-analyses encompassing 1671studies. The overall mean-weighted effect size (*d*) for reading methods was positive for all studies and moderate ($d$ = .43) across all studies. However, the researchers could not rule out the potential of publication bias and therefore are unable to draw firm conclusions from the study.

*Keywords:* meta-analysis, reading, student achievement

The Effects of Reading Methods on Student Achievement

The approach to reading instruction and the selection of strategies that develop proficient readers is at the center of major debates among educators (Hattie, 2009, p. 129). Central to the competing philosophies is the phonics versus whole language controversy and the numerous programs based on these approaches. Hattie (2009) synthesized more than 800 meta-analyses to determine what factors improve student achievement. Reading is the foundation for all other learning, therefore an important part of his work involves the effects of various reading protocols on student achievement. The purpose of this paper is to reconfigure the constructs utilized by Hattie that related to reading philosophies and methods in order to create a new category we identify as the effects of reading methods on student achievement. In order to deepen the thinking about different reading philosophies, we included the strategies of adjunct aids, a more visual approach to instruction, and direct instruction, which focuses solely on teacher learning objectives (Adams & Englemann, 1996), to Hattie's original analysis of distinct reading strategies.

## Method

Six constructs from the meta-analyses in the chapter on the contributions from the curricula (i.e., vocabulary programs, phonics instruction, sentence combining, repeated reading, comprehension programs, and whole language), and two constructs from the contributions from teaching approaches (part II) (i.e., adjunct aids and direct instruction), were synthesized to create a new higher-order meta-analysis on the effects of reading methods on student achievement. As shown in Table 1, all variables were analyzed.

[Insert Table 1 about here.]

The forest plot presented in Figure 1 was derived from pre-calculated effects using data from Hattie (2009, pp. 130, 201). This included all calculations done by Stata/IC version 12.1 comprised a table of effect sizes and their 95% confidence intervals (CI). Statistical heterogeneity was assessed using the $I^2$ statistic, where a value greater than 75% was deemed to indicate a high degree of heterogeneity and to require a random effects meta-analysis.

[Insert Figure 1 about here.]

**Data Analysis**

All studies were originally reported by Hattie (2009) and then the resultant findings were checked for this study against the original research articles. Coding was not necessary because the effect sizes were provided in Hattie. To determine the effects of reading methods on achievement, a data set was created that included the name of each variable, number of studies, common language effect (CLE), effect size, and the standard error (*SE*). For a graphic display of effect sizes and the 95% confidence intervals (CI) a forest plot was created using Stata/IC version 12.1 from pre-calculated effects using data from Hattie (2009, pp.130, 201). Utilizing the new data set, a funnel plot was created to test for publication bias.

## Results

Hattie's (2009) meta-analysis involved 46 meta-analyses, based on 1,671 studies summarizing the use of different reading methods encompassing students of all ages. The data set was developed using eight constructs presented in Hattie: repeated reading ($d$ = 0.67), vocabulary programs ($d$ = 0.67), phonics instruction ($d$ = 0.60), direct instruction ($d$ = 0.59), comprehension programs ($d$ = 0.58), adjunct aids ($d$ = 0.37), sentence combining ($d$ = 0.15), whole language ($d$ = 0.06), resulting in newly synthesized information on the effects of reading methods on achievement.

*Phonemic awareness.* According to the National Reading Panel (Langenberg et al., 2000) many strategies are used to improve phonemic awareness; phoneme isolation, identification, categorization, blending, segmentation, and deletion. "The National Reading Panel found an overall effect size on phonological outcomes of $d = 0.86$, reading outcomes of $d = 0.53$, and on spelling of $d = 0.59$." Fourteen meta-analyses on phonics instruction were conducted from 425 studies depicting an effect size of $d = 0.60$ with an $SE = 0.22$. The CLE for phonics instruction uncovered that if students were randomly selected from the treatment group, phonics instruction would show a positive effect 66% of the time, 16% greater than chance. Hattie (2009) determined the meta-analyses on phonics instruction is substantial in improving reading skills.

*Whole language*. The effect of whole language programs on reading achievement included four meta-analyses, which included 64 studies. Whole language programs had a low effect ($d = 0.06$) and an $SE$ of 0.06. Hattie (2009) concluded that whole language programs had negative effects on learning to read because reading strategies need to be explicitly taught, especially to students who struggle in reading. In a study by Jeynes and Littell (2000), students in the lower elementary grades with low socioeconomic status, performed better than their peers when using a basal reader. According to the CLE, if chosen at random from the treatment group, students who are instructed with whole language programs would have a 52% probability of having a desirable outcome on student achievement measures, only slightly higher than by chance. Because the confidence interval includes zero, we cannot be conclusive about whether this effect is statistically different from zero.

*Vocabulary programs*. Vocabulary programs had a greater effect than either whole language or phonics programs. The synthesis of seven meta-analyses that included 301 studies

on vocabulary programs had an overall effect size of $d = 0.67$ and an *SE* of 0.11, with the greatest mean effect size of $d = 0.97$ for those students who received direct vocabulary instruction as measured by "reading comprehension of passages containing taught words (Stahl & Fairbanks, 1986)." Determining word meaning from context ($d = 0.43$) and understanding the derivation of words also had positive effects on reading achievement (Fukkink & de Glopper, 1998). According to the CLE, if chosen at random from the treatment group, students who are exposed to vocabulary programs would have a 68% probability of having a more desirable outcome on student achievement measures.

*Repeated reading*. Repeated reading programs and its effect on reading achievement were determined by two meta-analyses that included 54 studies, a high effect size ($d = 0.67$), and an *SE* of 0.08. According to Chard, Vaughn, and Tyler (2002) repeated reading is defined as "re-reading a short and meaningful passage until a satisfactory level of fluency is reached." These researchers noted that the format of a test had the greatest effects in determining the variance between timed and untimed tests. The variance was due to students' abilities to read fluently. The CLE determined that students selected at random would have a 68% probability of having a desirable outcome on student achievement measures. Because the confidence interval includes zero, we cannot be conclusive that this effect is different from zero.

*Direct instruction*. Four meta-analyses including 304 studies on the use of direct instruction had a medium effect size ($d = 0.59$) on reading achievement with an *SE* of 0.10. Adams and Englemann (1996) defined direct instruction as having clearly stated goals that the teacher models with exemplars, facilitates guided practice with feedback, and follows up with independent practice in order to drive future instruction to ensure mastery. Hattie (2009) concurred that direct instruction provides students with opportunities to acquire new knowledge

and apply in different contexts. The CLE of direct instruction indicates that if selected at random from the treatment group, direct instruction had a 66% probability of having a desirable outcome on student achievement measures.

*Comprehension programs*. Hattie (2009) found comprehension programs with a medium effect size of $d = 0.58$ and an *SE* of 0.06 based on nine meta-analyses including 415 studies. This overall effect of comprehension programs included studies which place an emphasis on learning words (Swanborn & de Glopper, 2002) as well as higher order comprehension strategies. This included inferring, asking questions to set a purpose for reading, monitoring one's own comprehension during reading, and summarizing. The study by Guthrie, McRae, and Klauda (2007) found a high effect size ($d = 0.93$) on multiple text comprehension when using a workshop model (modeling, scaffolding, and guided practice). The CLE of comprehension programs found that if students were randomly selected from the treatment group, comprehension programs would have a 66% probability of having a desirable outcome on student achievement measures.

*Adjunct aids*. Four meta-analyses including 73 studies on using adjunct aids in texts had a low effect size ($d = 0.37$) on student achievement (Hattie, 2009). Hoeffler and Leutner (2007), have described adjunct aids as illustrations within texts that can be animated, decorative, or representational. The mean effect size of $d = 0.37$ with $SE = 0.04$, having a narrow range of 0.29 to 0.45 produces a large weight (16.16), so adjunct aids are the most dominant variable in this pooled effect size from this meta-analysis. According to the determined common language effect (CLE), a subject selected randomly from the treatment group provided with adjunct aids would have a 60% probability of having a desirable outcome on student achievement measures.

*Sentence combining*. The effect of sentence combining was determined by two meta-analyses that included 35 studies with a mean effect size of $d = 0.15$ and $SE = 0.09$. Sentence combining is operationalized as putting sentences together to convey a complex thought (Fusaro, 1993). A relatively large weight of 14.54, determined by a range of -0.02 to 0.32 results in this effect having substantial influence in the model. The CLE of sentence combining indicates that if selected at random from the treatment group, combining sentences had a 54% probability of having a desirable outcome on student achievement measures. Because the confidence interval includes zero, we cannot be conclusive about the effect of this variable.

*Reading methods*. When all of these effects were combined to create a new category, reading methods, an overall effect size of $d = 0.42$ was found for effects related to the effects of reading methods on achievement and was highly significant $p < .001$. Most of the studies chosen in the meta-analysis, except for adjunct aids ($d = 0.37$), sentence combining ($d = 0.15$), and whole language ($d = 0.06$), had effect sizes that met the criteria outlined by Hattie (2009) of minimum $d = .40$. To determine the probability that there is a 66% confidence that the effect will result in reading achievement, a forest plot was created using a minimum effect size of $d = 0.58$, as depicted in Figure 1. It is evident that five of the eight effects (i.e., repeated reading, vocabulary programs, phonics instruction, comprehension programs, and direct instruction) are most likely to have a positive impact on reading achievement.

Funnel plots are used to determine the potential for publication bias (Lipsey & Wilson, 2001, p. 143). As evidenced by the funnel plot depicted in Figure 2, the studies in the meta-analysis reflect potential publication bias. The visual evidence indicated an asymmetrical pattern. Repeated reading, which appears on the bottom of the funnel plot, was the outlier with a

wide CI [- 0.64, 1.98]. The next lowest effect size on the funnel plot was phonics instruction with a CI [0.17, 1.03]. The top study, adjunct aids, is reflective of the mean effect size.

[Insert Figure 2 about here.]

The weighted effects in the forest plot in Figure 1 show the impact of the constructs on the overall mean effect size. The construct, adjunct aids had the largest weight at 16.16% with the most precise CI [0.29, 0.45], even though it had a moderately low effect size of $d = 0.37$. Comprehension programs had the next highest weight at 15.76% with a CI of [0.47, 0.69]. Whole language also had a weight of 15.76%, even though it had the lowest effect size of $d = 0.06$ with a 95% CI [- 0.05, 0.17]. Sentence combining weighted next at 14.54% with a 95% CI of [- 0.02, 0.32], with a very low effect size of $d = 0.15$. Direct instruction followed with a weight of 14.13% and 95% CI [0.40, 0.78]. Vocabulary programs weighted at 13.56% with a 95% CI [0.46, 0.88]. The two variables with the lowest weights are phonics instruction with a weight of 8.43% and a 95% CI [0.17, 1.03] and lastly repeated reading had the lowest weight of 1.66% with a 95% CI [- 0.64, 1.98]. However, repeated reading had one of the highest effect sizes in our meta-analysis of $d = 0.67$.

The most influential weight on the overall construct of reading methods of 16.16% was from the variable adjunct aids, which had a moderately low effect of $d = 0.37$ and ranked 72$^{nd}$ overall in impact among the studies analyzed in Hattie (2009). Another noticeable result was the high weight of whole language programs, which had a very low effect of $d = 0.06$ and a high weight of 15.76%. Whole language programs ranked 129$^{th}$ overall in impact (Hattie, 2009). When all of the effects of reading methods were pooled there was a high degree of variation in adjunct aids and whole language. The effects of reading methods on achievement were pooled the $I^2$ value of 89.3% reveals a large amount of heterogeneity that was highly significant, $p <$

.001. Using the random effects analysis an overall effect size of $d = 0.42$ was determined. Based on the effects of reading methods on achievement, adjunct aids and reading comprehension programs have the greatest weight in the overall effect of reading method on student achievement.

## Discussion

Hattie (2009) did not report an oveall effect size for reading. The reading methods explored in this meta-analysis produced a small pooled effect size ($d = 0.43$). From this we conclude that purposeful approaches to reading instruction will impact learning. The debate between whole language and phonics methods distracts from research-based decisions educators need to make as they design reading programs. Based on our analysis, there is strong evidence to suggest that vocabulary programs, direct instruction and comprehension programs have greater impact on student achievement than the other approaches, including both phonics and whole language Hattie's conclusion that an eccelectic approach, that combines methods, will be the most effective is supported by the reconfigured construct of reading methods (p. 140). However, we are unable to draw a firm conclusion about the impact of these reading programs. The potential for publication bias, as indicated by the funnel plot, means that we cannot be confident that conclusions drawn from the data are valid.

References

*Adams, G. L. & Engelmann, S. (1996). *Research on direct instruction: 20 years beyond DISTAR.* Seattle, WA: Educational Achievement Systems.

*Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities, 35,* 386-406.

*Fukkink, R. G., & de Glopper, K. (1998). Effect of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68,* 450-469.

*Fusaro, J. A. (1993). A meta-analysis of the effect of sentence-combining on reading comprehension. *Reading Improvement, 30,* 228-231.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

*Hoeffler, T., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17,* 722-738.

*Langenberg, N. L., Correro, G., Ferguson, G., Kamil, M. L., & Shaywitz, S. E. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: National Institute of Child Health and Development.

Lipsey, M.W & Wilson, D.B. *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

*Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56,* 72-110.

Table 1

*Effects of Reading Methods on Achievement*

| Study | NumStudies | EffectSize | SE | CLE |
|---|---|---|---|---|
| AdjunctAids | 73 | 0.37 | 0.043 | 60% |
| SentenceCombining | 35 | 0.15 | 0.087 | 54% |
| RepeatedReading | 54 | 0.67 | 0.67 | 68% |
| VocabularyPrograms | 301 | 0.67 | 0.108 | 68% |
| PhonicsInstruction | 425 | 0.6 | 0.221 | 66% |
| ComprehensionPrograms | 415 | 0.58 | 0.056 | 66% |
| WholeLanguage | 64 | 0.06 | 0.056 | 52% |
| DirectInstruction | 304 | 0.59 | 0.096 | 66% |

*Note*. NumStudies = number of studies; EffectSize = effect size; CLE = Common Language Effect.

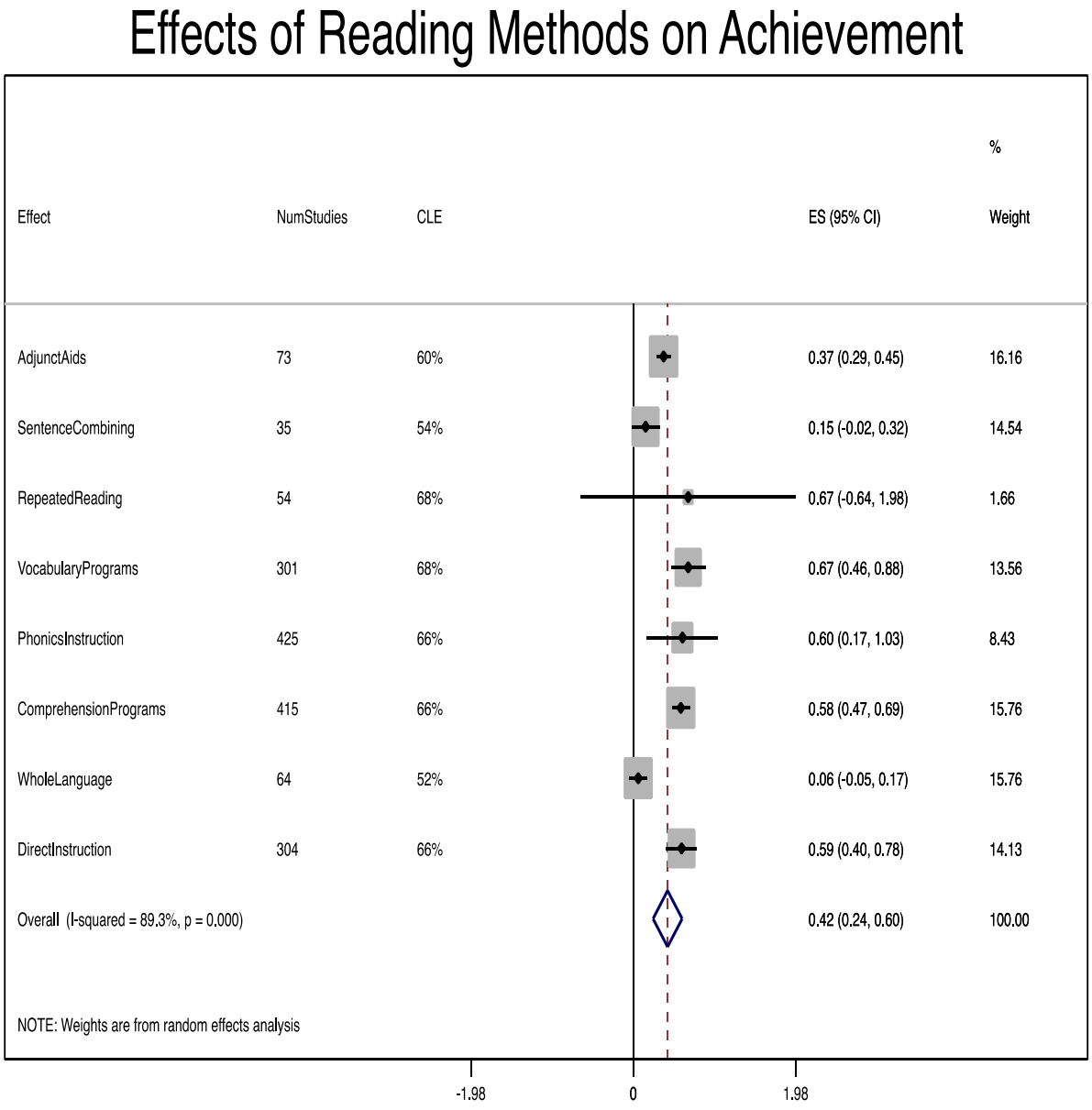



*Figure 1*. Forest plot predicting effects of reading methods on achievement. CLE = common language effect; ES = effect size; CI = confidence interval; $I^2$ = 89.3%; $p < .001$. Adapted from "Visible Learning," by J. C. Hattie, 2009, pp. 130, 201.

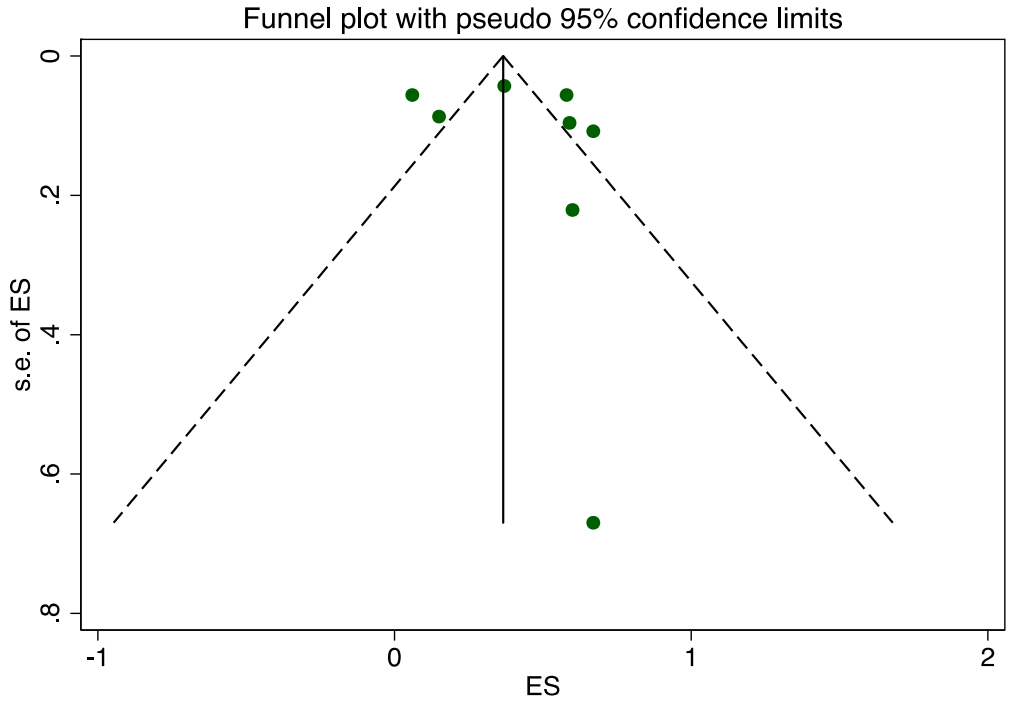



*Figure 2*. Funnel plot, using data from eight meta-analyses on the effects of reading methods on achievement.