Comprehension Meta-Analysis Update:

In late 2022, I completed a meta-analysis with Joshua King and Sky McGlynn on reading comprehension. This meta-analysis re-analyzed studies synthesized by two previous meta-analyses conducted by Filderman and colleagues (2022) and Silverman and colleagues (2020). We also screened 394 additional studies published after these earlier analyses. That said, our analysis was different from the original meta-analyses, as our inclusion criteria was more restrictive. While the previous meta-analyses included assessments not directly related to reading comprehension, we did not. Thus, the prior meta-analyses provided a broader estimate of intervention impacts but with less specificity regarding reading comprehension outcomes. In total, we meta-analyzed 80 studies.

We conducted this meta-analysis because we believed there was a critical, unresolved question regarding reading comprehension. Previous meta-analyses on background knowledge and vocabulary instruction, such as those by Elleman and colleagues (2009), Silverman et al. (2020), and Hwang and Cabell (2022), suggested that content and vocabulary instruction showed large effect sizes on researcher-designed assessments, but significantly smaller effects on standardized assessments. This suggested that while vocabulary and content instruction support comprehension in texts directly related to the instruction, this benefit does not necessarily transfer to general reading comprehension. The comprehensive Filderman et al. (2022) meta-analysis indicated that strategy instruction significantly impacted comprehension outcomes, yet it did not control for assessment type. Therefore, the question remained: Does strategy instruction produce better transfer effects than content and vocabulary instruction, or does it also show weaker results on standardized reading comprehension tests?

We submitted the full article for peer review, and I published a summary on this blog. The post became one of my most controversial, likely because our findings challenged a popular belief of the Science of Reading movement. Our analysis suggested that background knowledge and vocabulary contributed the most to reading comprehension improvement, but we also found this effect did not transfer to new texts unrelated to the background knowledge or vocabulary instruction. In contrast, strategy instruction demonstrated a statistically significant positive transfer effect, challenging the prevailing view that background knowledge outweighs strategy instruction in importance for developing reading comprehension ability.

Since then, our meta-analysis has undergone two full rounds of peer review. First, we submitted to the *Review of Educational Research,* where reviewers cited four main criticisms:

- Lack of detail in the introduction and methods sections.
- Insufficient breadth in the systematic search, (though this criticism surprised me given that it was -to the best of my knowledge- the largest meta-analysis on the topic.)
- Coding was performed by only one of the three authors (myself).

A year later, I updated the meta-analysis write up to address some of these critiques and resubmitted it to *Reading Research Quarterly.* However, the second review yielded similar criticisms, especially the concern about coding by only one author. After a pause, I revised the manuscript again this summer with support from Dr. Kathryn Garforth, addressing reviewer feedback, updating the analysis with new studies published since the original manuscript, and improving the meta-analysis based on my own research development. This led to several significant updates:

1. An additional 1,060 studies were screened to ensure the meta-analysis remained current through 2024.

2. We included all studies from the Hwang and Cabell (2022) meta-analysis. (I believe it was an oversight to have not done this in the first place, as it was the largest peer-reviewed meta-analysis of content instruction during literacy lessons.

3. One-quarter of the studies were dual-coded by Dr. Kathryn Garforth to improve validity.

4. Error rates were estimated for increased transparency.

5. Effect size calculators and coding databases were uploaded to publicly accessible servers to adhere to open science principles.

6. Studies lacking sufficient statistical detail for independent effect size and standard error calculations were excluded.

7. Outliers were trimmed using Z-scores and an IQR analysis, resulting in more homogeneous results, especially for weighted effect sizes of proximal assessments. Consequently, this version yielded a much lower mean effect size than the original manuscript.

8. Effect size weights were updated to reflect pooled inverse variances between treatment and control groups, rather than the mean variances.

9. Weights were recalculated using post-test sample sizes for individual assessments, instead of pre-attrition sample sizes.

10. Three additional regression analyses were conducted, examining combinations of pedagogies, instructional dose, and estimated fixed effects of pedagogies.

11. The entire manuscript was rewritten to better incorporate peer-review feedback.

These changes aim to enhance the peer-review process and increase the transparency and validity of our findings. Reflecting on this project, one of my earliest attempts at publishing peer-reviewed research, I can say that while it has been more challenging than anticipated, it has provided invaluable learning.

To update my readers, as the current website no longer reflects my most rigorous analysis, I have uploaded the current working manuscript to the OSF pre-print database. By sharing it on the OSF, I hope to receive further feedback before resubmitting it to a peer-reviewed journal. Over the next few months, there will be additional revisions from my co-authors as we prepare for a third submission.

Despite these methodological updates, our core findings remained largely consistent. We found a greater transfer effect for strategy instruction than for knowledge and vocabulary instruction. Knowledge and vocabulary instruction maintained a significant impact on researcher-designed assessments but not on standardized assessments. Strategy instruction's impact diminished over time, whereas knowledge instruction's impact improved. Reciprocal teaching emerged as the most effective intervention for reading comprehension outcomes.

However, three new findings emerged:

- Knowledge instruction led to statistically significant improvements on standardized reading comprehension assessments when studies were conducted in core instructional settings for students in grade 3 and above.
- The highest overall effect sizes were observed in studies combining vocabulary, background knowledge, and cognitive strategies.
- Only cognitive strategies and reciprocal strategies consistently improved reading comprehension outcomes. Other strategies had negligible or even negative impacts on average.

These new insights suggest that my initial stance—that knowledge instruction did not enhance generalized reading comprehension—may require more nuance. Below is a more in-depth summary of the meta-analysis results, and the full pre-print manuscript is available at the following link:

2024 Reading Comprehension Meta-Analysis: Summary

A total of 1,557 studies were screened from the Education Source and ERIC databases, and three previous meta-analyses on the topic: Hwang & Cabell (2022), Filderman et al. (2022), and Silverman et al. (2020). To qualify, studies had to be experimental or quasi-experimental, examine the impact of a reading comprehension pedagogy, include a reading comprehension assessment, have a minimum sample size of 20, and report enough information to calculate both a pooled standard error and a Cohen's d effect size. We evaluated 92 studies and 141 mean effect sizes.

On standardized tests, mean effect sizes were as follows: vocabulary (.11), content (.17), cognitive strategies (.21), metacognitive strategies (.06), reciprocal teaching (.45), morphology (.13), and graphic organizers (.15). Only cognitive strategies and reciprocal teaching produced a statistically significant transfer effect, as can be seen in Figure 1.

Figure 1. *Standardized Test Results*

The Impact of Reading Comprehension Pedagogies on Standardized Comprehension Tests.
Hansford, Garforth, McGlynn, and King (2024) [Pre Print]
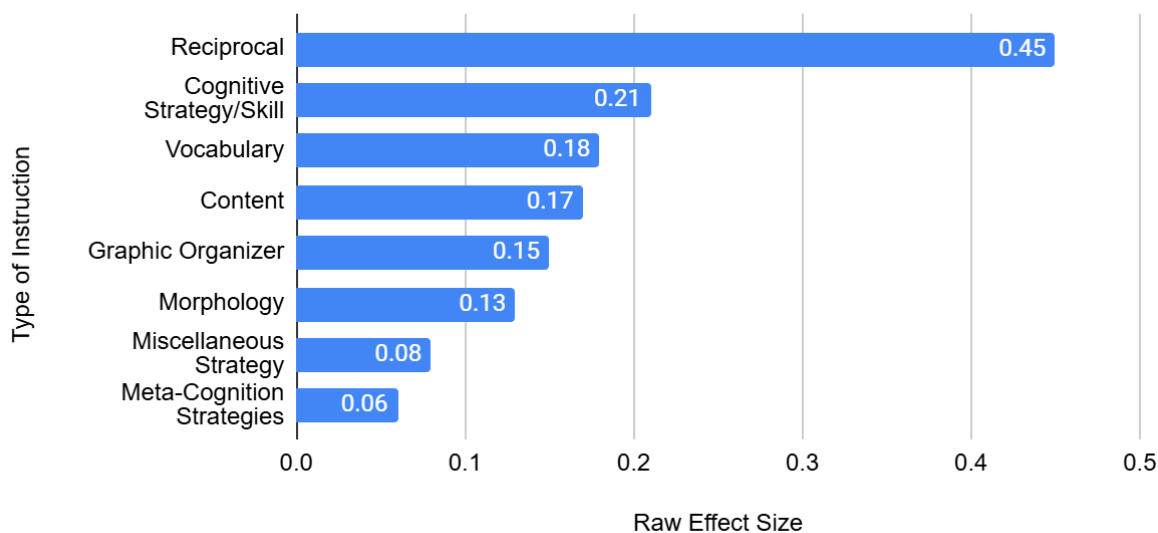
While the results of Figure 1, help to estimate what comprehension pedagogies best improve comprehension outcomes, all these effect sizes are contextually dependent. Nineteen additional moderator and regression analyses indicated that effect sizes depended on factors such as student age, academic need, pedagogical combinations, duration, and research quality. Although the primary goal was to determine which pedagogies best transferred to distal reading comprehension, the findings were very context dependent.

For core settings, weighted effect sizes were significant on standardized reading assessments for vocabulary (d = .34), content (d = .54), cognitive strategies (d = .26), and graphic organizers (d = .22). For intervention settings, only reciprocal teaching showed a significant effect size (d = .45). Figures 2, 3, show the standardized test results for core, and intervention settings. Figures were excluded for ELL settings, due to a small pool of studies.

Figure 2. *Core Results*

The Impact of Comprehension Pedagogies on Standardized Reading Comprehension Assessments, in Core Settings

Hansford, Garforth, McGlynn, and King (2024) [Pre Print]
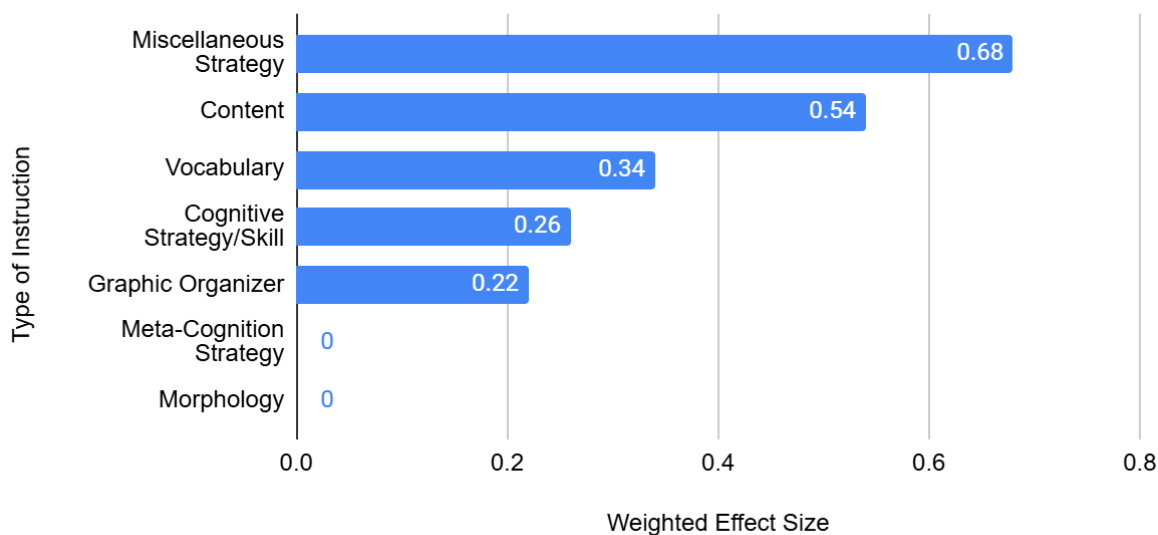
Figure 3. *Intervention Results*

The Impact of Comprehension Pedagogies on Standardized Reading Comprehension Assessments, in Intervention Settings

Hansford, Garforth, McGlynn, and King (2024) [Pre Print]
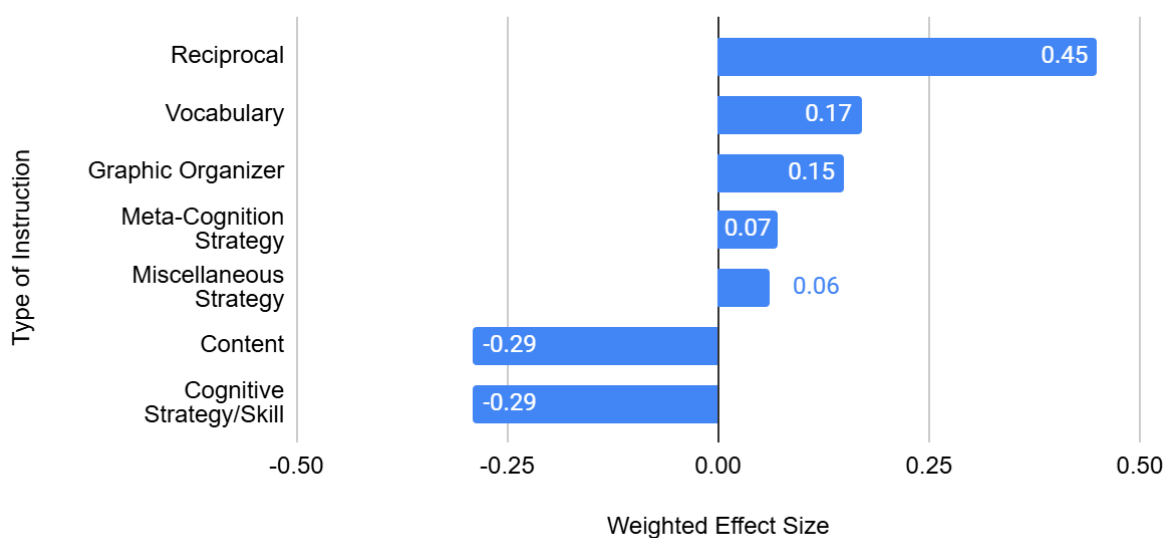
Among kindergarten studies, cognitive strategies and vocabulary were significant (d = .23 and .51, respectively). In primary grades, only vocabulary had a significant weighted effect (d = .28). In junior grades, graphic organizers were significant (d = .47), and in intermediate grades, meta-cognitive strategies were significant (d = .43). For secondary grades, only reciprocal teaching showed a significant weighted effect (d = .89), although based on a single study. Another secondary study showed a large raw effect for vocabulary, content, and cognitive strategies (.46), but due to a small sample size, the weighted effect size was low (d = .08). Figures 4, 5, and 6 show the results for grades 1-8. Figures were excluded for kindergarten and secondary school, due to a small pool of studies.

Figure 4. *Grades 1-3*

The Impact of Comprehension Pedagogies on Standardized Reading Comprehension Assessments, for Grades 1-3

Hansford, Garforth, McGlynn, and King (2024) [Pre Print]

Figure 5. *Grades 4-6*

The Impact of Comprehension Pedagogies on Standardized Reading Comprehension Assessments, for Grades 4-6

Hansford, Garforth, McGlynn, and King (2024) [Pre Print]
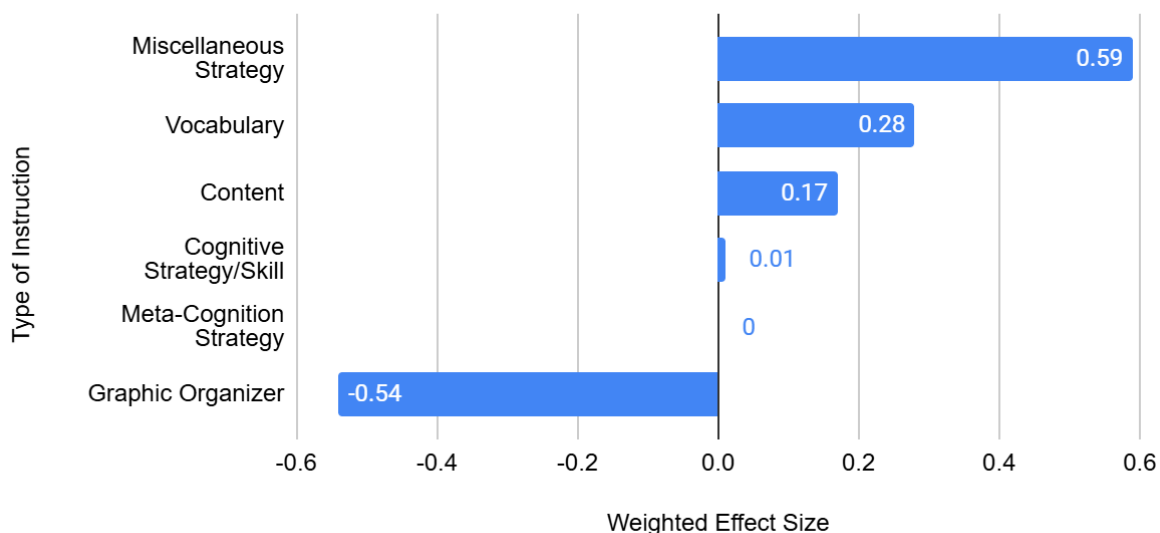

Figure 6. *Grades 7-8*

The Impact of Comprehension Pedagogies on Standardized Reading Comprehension Assessments, in Grades 7-8

Hansford, Garforth, McGlynn, and King (2024) [Pre Print]
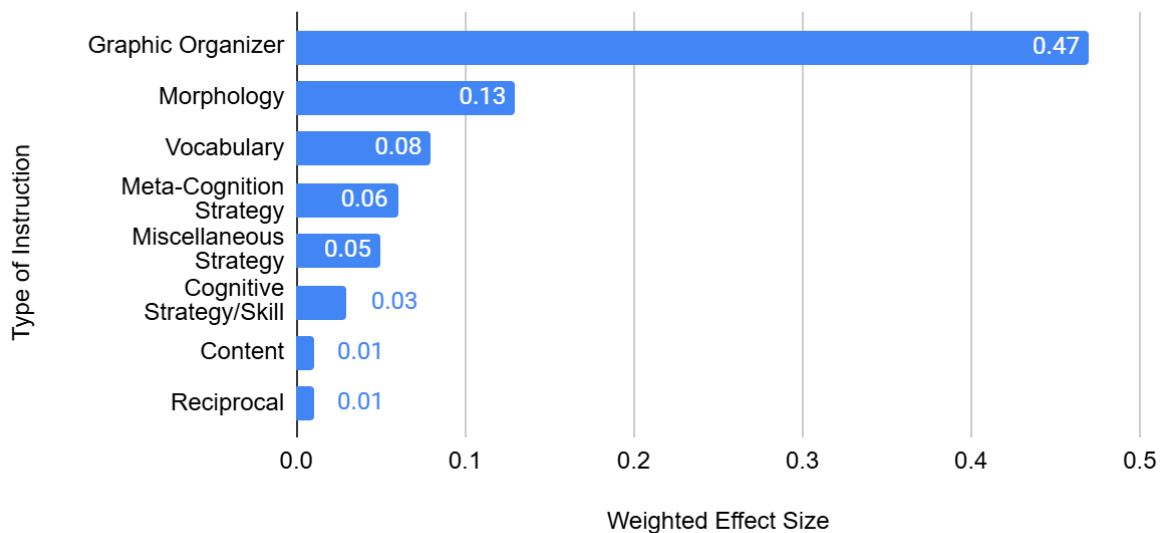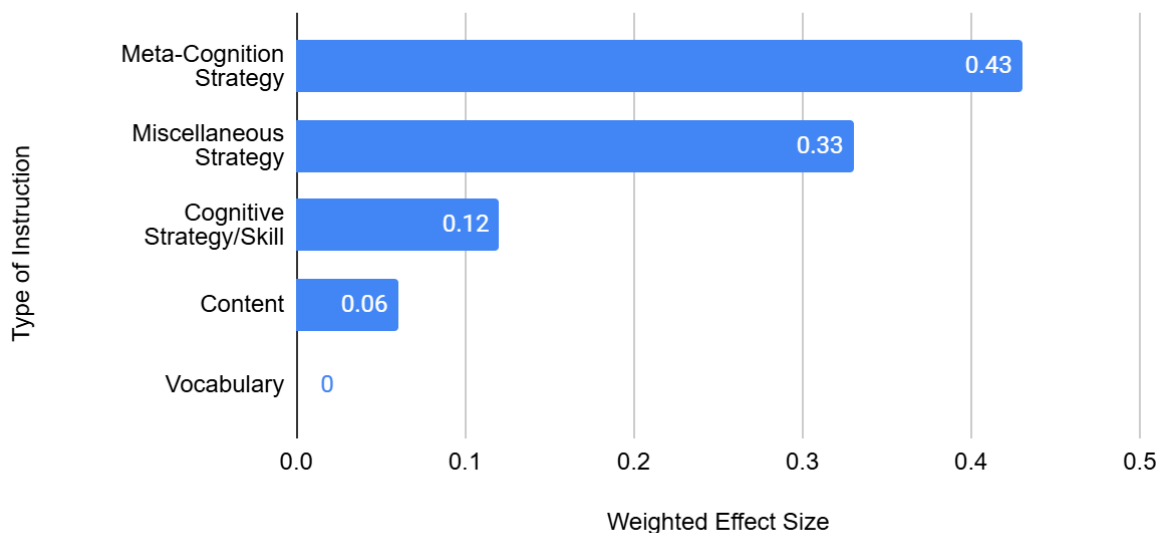
A regression analysis comparing pedagogy combinations revealed the highest comprehension results occurred when content, vocabulary, and cognitive strategies were combined, with a mean ES of .51. Vocabulary combined with any pedagogy had a synergistic effect. When paired with cognitive strategies, vocabulary had a mean weighted effect size of .44, and with content instruction, .50. Cognitive strategies made a substantial positive difference when included vs excluded (difference = .19). Figure 7. Shows the results of the regression analysis on combinations of comprehension pedagogies.

Figure 7. *Regression Analysis*

Comparing Different Combinations of Reading Comprehension Pedagogies

Hansford, Garforth, McGlynn, and King (2024) [Pre Print]
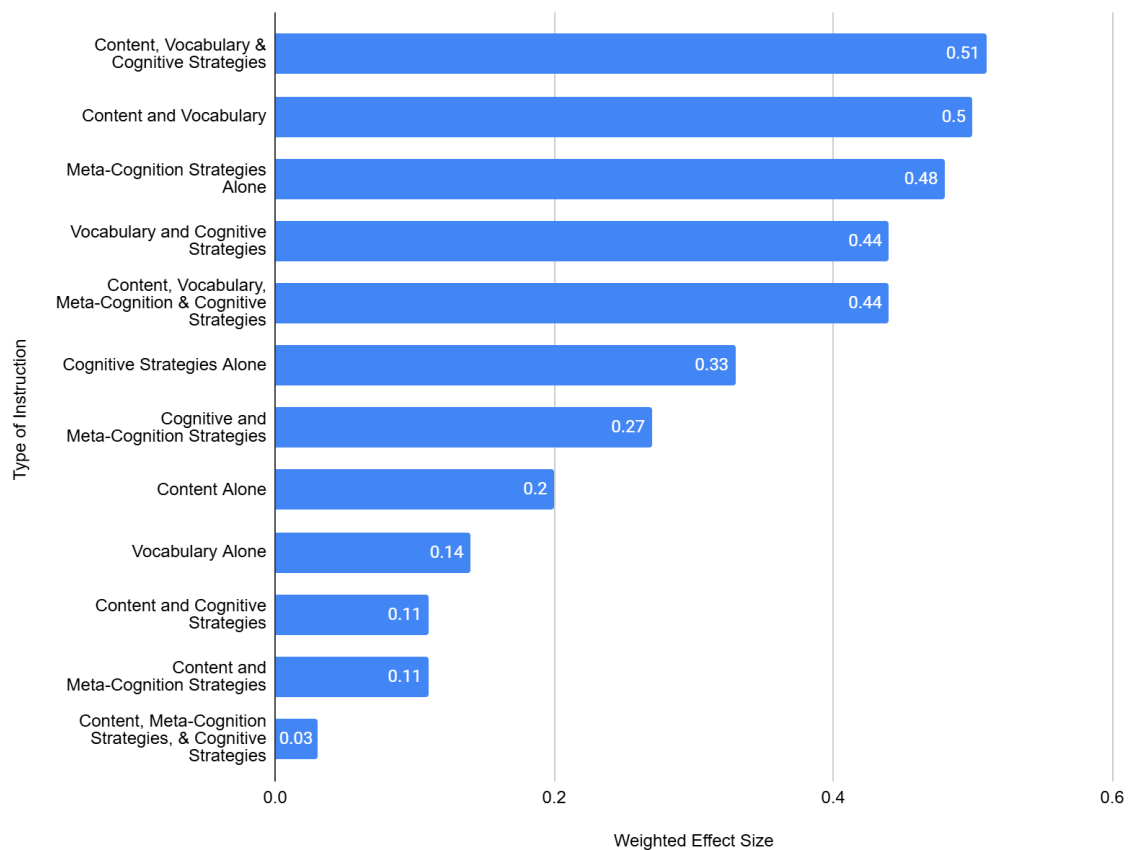
This meta-analysis suggests favouring nonfiction for literacy instruction. However, a neurological study by Aboud et al. (2019) indicated different brain areas activate for fiction and nonfiction, suggesting a potential need to develop comprehension skills for both text types, though the idea of comprehension as a teachable skill remains contested (Willingham, 2009).

Research rigor was controlled via regression analysis (see Tables 17, 18, and 19). This analysis indicated that most studies included multiple experimental variables, non-randomized samples, proximal assessments, or small sample sizes. Only a few studies used standardized distal measurements, randomized samples, and controlled for a single experimental variable. Although these variables reflect authentic learning environments rather than research limitations, they complicate causal inference. Content and vocabulary instruction studies tended to be more rigorous, but more high-quality studies are still needed, as are studies comparing different pedagogy combinations and durations.

To simplify these findings for teachers, I've created an infographic, in figures 8-9. Under each heading, pedagogies with evidence of efficacy are listed. Evidence of efficacy was defined as either a mean effect size or a weighted mean effect size of .20 or higher, on standardized reading assessments. I do want to note that the absence of a pedagogy does not imply ineffectiveness but rather a lack of supporting evidence from this analysis. Pedagogies are listed by weighted effect size.

Figure 8. *Teacher recommendations*

Teacher Recommendations

Instructional Setting

Pedagogies Found to Be Effective

Core
- comprehension strategies
- background knowledge
- vocabulary
- cognitive strategies
- graphic organizers

Intervention
- reciprocal Teaching
- vocabulary
- meta-cognition strategies

ELL
- reciprocal Teaching
- comprehension Strategies
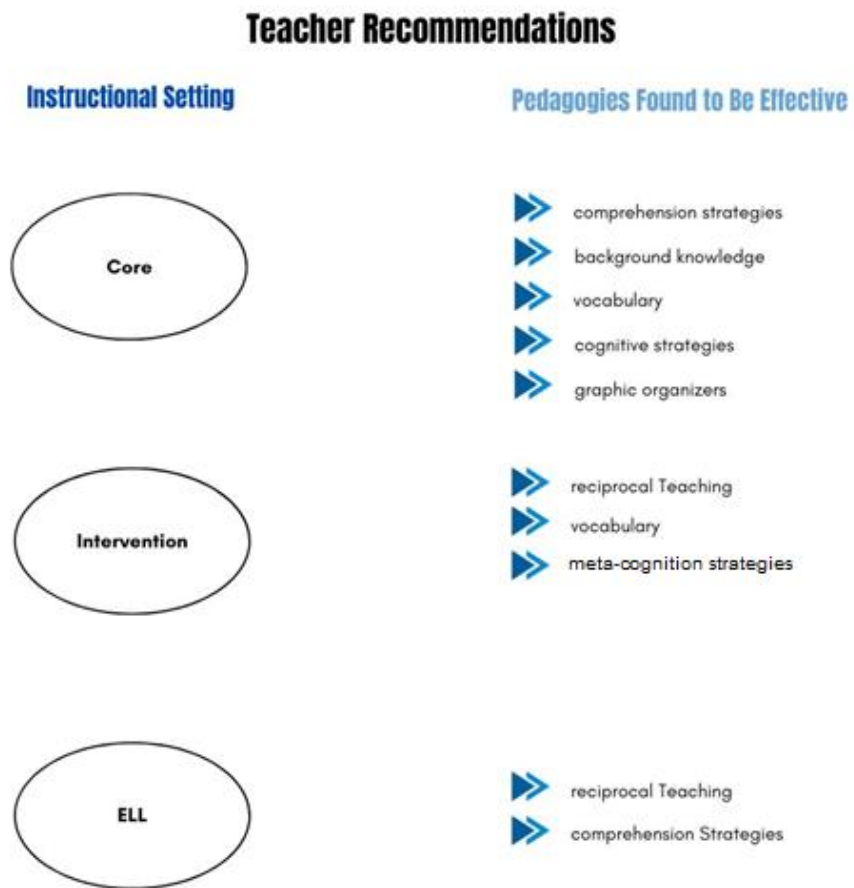
Figure 9. *Teacher recommendations continued*

Instructional Setting

Pre-K / Kindergarten

Grades 1-3

Grades 4-6

Grades 7-8

Grades 9-12

Pedagogies Found to Be Effective

» vocabulary
» cognitive strategies

» comprehension strategies
» vocabulary
» cognitive strategies

» graphic organizers
» reciprocal teaching

» meta-cognition strategies
» comprehension strategies

» reciprocal teaching
» vocabulary
» background knowledge
» cognitive strategies

One final note: this meta-analysis is nearly 100 pages long, with the results section alone spanning 36 pages. We analyzed the data in every conceivable way to leave as few unanswered questions as possible and to ensure maximum transparency. However, by doing so, I am concerned that it may create opportunities for cherry-picking. With such a wide range of moderator and regression analyses conducted, there is the risk that individuals might

selectively use the data to support pre-existing biases rather than forming opinions based on the data itself.

With this in mind, I encourage readers to adopt a nuanced approach. This meta-analysis provides at times contradictory evidence. However, overall, it suggests robust support for vocabulary, content, cognitive strategies, and reciprocal teaching. It does not, however, strongly support metacognitive strategies or miscellaneous comprehension strategies. Additionally, this analysis does not offer substantial empirical backing for the necessity of a knowledge-building curriculum. We concur with Hwang and Cabell (2022) that, at present, there is insufficient research to fully evaluate this question via meta-analysis. Nonetheless, I believe this meta-analysis strongly supports an increased focus on non-fiction instruction over fiction during the literacy block, especially for grades 3 through 12.

Written by Nathaniel Hansford

Last Updated: 2024/11/18

References:

-Elleman, A.M., Lindo, E.J., Morphy, P., & Compton, D.L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness,* 2 (1), 1 – 44. https://doi-org.ezproxy.lakeheadu.ca/10.1080/19345740802539200

-Burns, Matthew & Duke, Nell & Cartwright, Kelly. (2023). Evaluating components of the active view of reading as intervention targets: Implications for social justice. *School Psychology* (Washington, D.C.). 38. 30-41. 10.1037/spq0000519.

-Filderman, M. J., Austin, C. R., Boucher, A. N., O'Donnell, K., & Swanson, E. A. (2022). A Meta-Analysis of the Effects of Reading Comprehension Interventions on the Reading Comprehension Outcomes of Struggling Readers in Third Through 12th Grades. Exceptional Children, 88(2), 163-184. https://doi.org/10.1177/00144029211050860

-Hwang, H., Cabell, S. Q., & Joyner, R. E. (2022). Effects of integrated literacy and content-area instruction on vocabulary and comprehension in the elementary years: A meta-analysis. *Scientific Studies of Reading, 26*(3), 223–249. https://doi.org/10.1080/10888438.2021.1954005

-Peng, P., Wang, W., Filderman, M. J., Zhang, W., & Lin, L. (2024). The active ingredient in reading comprehension strategy intervention for struggling readers: A Bayesian network meta-analysis. *Review of Educational Research, 94*(2), 228–267. https://doi.org/10.3102/00346543231171345

-Willingham, D. (2006). The usefulness of brief instruction in reading comprehension strategies. Ask the Cognitive Scientist [Blog]. Accessed at https://education.ufl.edu/patterson/files/2020/10/Willingham-ComprehensionStrategies.pdf on November 14th, 2024.