Article Summary:

Some reading researchers suggest that teachers should focus more on teaching background knowledge and vocabulary and less on reading comprehension strategies. This is despite the fact that experimental research shows stronger outcomes for strategy instruction, than for background knowledge instruction, when standardized assessments are used. These researchers claim that the effect of strategy instruction diminishes over time, but that the effect of background knowledge instruction and vocabulary increases with time.

For the purposes of validating this claim, I examined 8 meta-analyses and 73 reading comprehension studies via both a random effects moderator analysis and a correlation analysis. I found moderate evidence that strategy instruction decreased in effectiveness over time, weak evidence that vocabulary instruction decreased in effectiveness over time, and weak evidence that content instruction increased in effectiveness over time.

Full disclosure, this article is far more technical than my average content. I therefore included a graphic below, which includes a summary of the effective range of dosage for different types of reading comprehension instruction, (effective being defined as effect sizes above .20). That said, I encourage readers to read further to understand how I came to these recommendations.

Instructional Type	**Dose**
Vocabulary	7-35 weeks
Cognitive Skills	6 weeks or less (per 1-2 strategies)
Meta-Cognition	2 -12 weeks (per 1-2 strategies)
Graphic Organizers	13 weeks or more
Reciprocal Teaching	7-35 weeks
Content Knowledge	1 Year or Longer

Table 11: *The Effective Range of Dose for Comprehension Pedagogies Per Year of Study*

Introduction:

Recently, I published a meta-analysis on reading comprehension instruction that compared the results of 9 different types of interventions on standardized and non-standardized assessments. The meta-analysis garnered a lot of attention, because the results for strategy instruction were on average higher than the impact of vocabulary and content knowledge instruction on standardized reading comprehension assessments.

In general, this seemed to be an unpopular finding, because many have hypothesized that content knowledge and vocabulary are the most important forms of instruction for improving reading comprehension. Indeed, I received a lot of questions and criticisms of the finding. Four criticisms seemed particularly persistent.

1. Standardized tests are not valid measurements, because they don't contain a lot of background knowledge. Because, by design, they minimize the need for background knowledge.
2. Strategies have been overused in the American school system and teachers are expected to spend too much time teaching cognitive strategies to young students.
3. There is a diminishing return to strategy instruction, but not for content or vocabulary instruction.
4. All explicit instruction includes knowledge; therefore, measuring the impact of knowledge instruction seems needlessly pedantic.

I wanted to take some time to address these criticisms/questions.

1. Are standardized tests less valid, because they don't contain a lot of background knowledge? Afterall the highest results in my meta-analysis for non-standardized assessments were for background knowledge instruction.

My biggest concern with using researcher designed assessments here, is that the effect sizes seem so inflated that I am not sure what they mean. On average studies that used their own custom assessments showed 4x higher results than studies that used previously created standardized assessments. These differences were especially pronounced for the impact of content instruction. On standardized tests, content instruction showed on average no meaningful benefit for reading comprehension; however, for custom assessments the results for content instruction were the highest.

This makes sense, as custom based assessments are often specifically tailored to the experimental instruction. For example, one study taught the treatment group about volcanoes, but not the control group. Both groups were tested on their comprehension by giving them a text about volcanoes. Of course, the group that had learned about volcanoes did better, but would that group have still done better if the test had been about snakes of the Amazon?

The results of my meta-analysis do not suggest that we should not teach background knowledge, but rather that we should make sure that the content we are teaching is aligned to what we want to use to test students' reading comprehension. That said, I do wonder: If the comprehension benefits found for background knowledge instruction only appear when the test is aligned with the instruction, are we actually testing for comprehension or are we testing for specific background knowledge?

2. Teachers are being excessively asked to teach cognitive strategies, such as finding the main idea, even in the primary grades. Isn't this a problem?

My research does not lend support for excessive practice with cognitive strategies, as I found by far the highest effect sizes for vocabulary instruction in the primary grades. As most texts in primary grades are simple narratives and do not contain abstract ideas, I see little value in main idea instruction during this time frame. The purpose of cognitive strategies, like "finding the main idea" is to help students find meaning in an abstract text. For this to be meaningful, the student has to be working with a text with abstract ideas. Moreover, as this article will show, no research suggests that main idea instruction should be taught for an entire year.

3. Does strategy instruction show diminishing returns over time, and does content/vocabulary instruction show increasing returns over time?

I did try to look at this claim in my original article by examining longitudinal studies. However, at the time of my meta-analysis, I found only one longitudinal study which used a standardized reading comprehension assessment, and it showed no meaningful benefit. (That said, two additional studies have since been published and I will review them at the end of this article.) Recently, A popular researcher cited a bibliography as proof for this claim, and I decided to dive deeper to get a better understanding of this question, starting with going through this bibliography (as can be seen below).

4. Is not all explicit teaching the instruction of knowledge? How could that be a bad thing?

Yes, all explicit instruction should be done for the purpose of building student knowledge. And no, building student knowledge is not a bad thing. However, this is not the debate at hand. The real question is do we need to use a scripted curriculum that systematically builds student knowledge over time, based on a predetermined scope and sequence as others have hypothesized. Or is such an approach unnecessary due to the impossibility of pre-determining what background knowledge is most important? Moreover, how does culturally relevant pedagogy factor into this discussion?

Study	Type of Study	Topic of Study	Number of Studies	Results
Donegan and Wanzek 2021	Meta-Analysis	Reading interventions for struggling readers	33 (case studies excluded)	-For intensity, no evidence that longer or individualized interventions were associated with larger effects. -Only comprehension interventions below 15 hours of instruction showed non-negligible effects. However, the p-value was greater than .05.
Fukkink & de Glopper 1988	Meta-Analysis	(Using context clues) -Related to 3-cueing	21 (case studies excluded)	"None of the methodological predictors reduced heterogeneity significantly"
Rosenshine 1996	Meta-Analysis	-Generating Questions/Reciprocal Teaching	26 (case studies excluded)	-No difference between 4 and 25 sessions for results
Berkeley 2009	Meta-Analysis	-Strategies and comprehension instruction	40 (case studies included)	-Researcher Designed Tests: 1 week or less d= .48 k=7 1-4 weeks d= .84 k=20 4 weeks or more d= .50 k=3 -Standardized Tests: 1-4 weeks: d= .49 k=7 4 weeks or more: d= .53 k=8

Table 1: *Previous Meta-Analyses that Examined the Dosage Effects for Comprehension Instruction*

Study	Type of Study	Topic of Study	Number of Studies	Results
Elbaum 2000	Meta-Analysis	1-1 reading interventions	29 (case studies included)	Not Applicable
Gajira 2007	Literature Review	Cognitive strategies and reciprocal teaching	29 (case studies included)	-Less than 4 hours of hours of intervention (mean ES = 1.61, SD = 0.91, n = 9), -Between 4 and 8 hours (mean ES = 1.34, SD = 0.58, n = 5), or -More than 8 hours (mean ES = 1.41, SD = 1.64, n = 8).
Talbot 1994	Literature Review	Comprehension Interventions (Vocabulary and Strategy)	48 (case studies excluded)	-The relationship between length of treatment and effect size was not significant (r=-.012, p<.90)
Suggate 2010	Meta-Analysis	Comprehension Interventions (not defined)	85 (case studies excluded)	-Comprehension Studies Mean d=.58 -Longitudinal Results d=.69

Table 2: *Previous Meta-Analyses that Examined the Dosage Effects for Comprehension Instruction, Continued*

The above studies have been cited to suggest that there is a decaying effect for teaching reading comprehension strategies. However, after reviewing each meta-analysis, I have several concerns with this precise interpretation.

These results seem to show no relationship, not a diminishing return. This relationship is especially problematic to the original claim if we consider that an effect size naturally corrects for time when measured against a control group. As effect sizes are measured by comparing the difference in learning between two groups at the end of a study's duration. A consistent effect would therefore suggest that the intervention remains to be effective, not that there is no additional benefit, by extending the intervention duration.

Second, most of these studies looked at multiple forms of comprehension instruction, including cognitive strategies, vocabulary, and reciprocal teaching, which makes it difficult to extrapolate whether these differences are specific to strategies or comprehension instruction in general.

Third, only one of these studies differentiated between this effect on standardized assessments vs non-standardized assessments, and non-standardized assessments show very inflated effects. The only study which did differentiate in this way showed higher results for longer strategy studies when they used a standardized test as the measurement.

Fourth, only one study showed a true diminishing effect (Donegan 2021). This study did not differentiate this effect based on standardized vs non-standardized assessments. It also included both vocabulary and strategy instruction under the same coding. The Donegan study was also based on a small number of studies and examined different factors.

Fifth, while looking at moderator variables gives us more specific information, this information is also inherently less well studied. As the more specific the research question becomes, fewer studies will be able to answer the question, and so the lower our confidences should be in the findings.

Sixth, using effect sizes to find diminishing returns for any intervention is difficult, as longer studies tend to show smaller effect sizes in general. For example, my meta-analysis on language programs showed much smaller effect sizes for longitudinal studies on both phonics programs and balanced literacy programs. This correlation might exist because study duration is a marker of study quality, and higher quality studies tend to show lower results. This correlation might also be caused by novel interventions being more stimulating or because instructors show an attrition in their fidelity over time.

Admittedly, comparing effect sizes to duration to try and find an ideal dosage is a common practice and an analysis I have also used many times. Specifically, this methodology has been used by many (myself included) to interpret the phonemic awareness results in the NRP meta-analysis. However, in retrospect, I think we should be very careful about drawing firm conclusions from this type of analysis.

Realistically, the above studies do not prove or disprove the hypothesis that strategy instruction benefits diminish over time, as the studies were not intended to answer this question. However, I decided to try and answer this question, now being mindful of the fact that effect sizes used in this way can be very misleading, I first conducted a correlation comparison between study duration and study effect sizes for each of the specific intervention types in my comprehension meta-analysis and compared the relative r effect sizes to the p-values. The results can be seen below.

Pearson Analysis for Standardized Assessment Results:

Instruction Type	Number of Effects	Pearson ES	P-Value
Vocabulary	18	-.31	>.001
Strategy	33	-.44	>.001
Reciprocal Teaching	7	-.32	>.001
Cognitive Strategy/Skill	26	-.30	>.001
Graphic Organizers	8	.40	>.001
Content/Background Knowledge	9	.30	>.001
Meta-Cognition Strategies	14	-.18	>.001
Technology	5	0	>.001
Morphology	4	-.28	>.001

Table 3: *Standardized Assessment Correlation Results*

Pearson Analysis for Custom Assessments

Instruction Type	Number of Effects	Pearson ES	P-Value
Vocabulary	20	.02	.0001
Strategy	22	.11	.0004
Reciprocal Teaching	4	.61	.0001
Cognitive Strategy/Skill	27	-.08	.0001
Graphic Organizers	12	-.15	.0002
Content/Background Knowledge	12	.82	.0002
Meta-Cognition Strategies	12	1.00	0
Technology	9	.35	.0001
Morphology	1	NA	NA

Table 4: *Correlation Effects on Researcher Designed Assessments*

If we examine the standardized test results, we see a clear trend. Time had a small negative impact on all interventions, except content knowledge and graphic organizers. Graphic organizers and content knowledge instruction seemed to improve student reading comprehension results, the longer the intervention was used. All p-values were insignificant. If we look at the custom assessment results, the P values are higher and the results less consistent. On custom assessments there were no statistically significant negative impacts for time. However, content knowledge instruction, technology use, meta-cognition strategies, and reciprocal teaching studies all showed higher effect sizes when studies were longer in duration.

One issue with this type of analysis is the fact that it was based on random effects. Most comprehension studies use multiple types of comprehension instruction at once. For example, it is not uncommon to see background knowledge, vocabulary, and strategy instruction in the

same study. In this model of analysis, we count such a study as a study on each of the three types of instruction. Similarly, most studies use business as usual control groups, in which the instruction in the control group is not controlled for. This means that the effect sizes found are more random and are therefore only meaningful if we have many studies or if we find very consistent effects.

In a fixed effect study, all the variables are the same between the control group and the treatment group, except for the variable being tested. For example, a fixed effect study might have the same content and instruction in both groups, but one group also receives vocabulary instruction. In this study design, we can be sure that the study results were being caused by the variable being tested. Unfortunately, most studies don't use this design, so relying on only a fixed effect analysis means relying on far less data to inform our understanding of science. That all said, when trying to answer a question as specific as "Does strategy instruction produce a diminishing return over time?" It might make sense to use a fixed effects model. So, I redid the Pearson analysis one more time, using only studies that measured a fixed effect and had a standardized assessment. The results can be seen below.

Fixed Effects Pearson Analysis

Instruction Type	Number of Effects	Pearson ES	P-Value
Vocabulary	1	NA	
Strategy	5	-.93	.07
Reciprocal Teaching	2	NA	NA
Cognitive Strategy/Skill	7	-.63	.04
Graphic Organizers	1	NA	NA
Content/Background Knowledge	1	NA	NA
Meta-Cognition Strategies	4	.57	.02
Technology	0	NA	NA
Morphology	1	NA	NA

Table 5: *Fixed Effects Pearson Analysis*

Unfortunately, only miscellaneous strategies, cognitive strategies, and meta-cognition strategies could be analyzed in this way. That said, miscellaneous strategies and cognitive strategies both showed a strong negative correlation, whereas meta-cognition strategies showed a strong positive correlation, meaning that longer studies showed lower effect sizes for miscellaneous strategies and cognitive strategies. Conversely, longer studies showed higher effect sizes for meta-cognition studies.

In general, I would argue all of these Pearson analyses provide moderate support for the idea that cognitive strategies, miscellaneous strategies, reciprocal teaching, vocabulary, and morphology instruction show diminishing returns over time. We also have moderate evidence that graphic organizers and content knowledge instruction show increasing returns over time. Lastly, we have mixed evidence on the impact of time on meta-cognition strategies' effectiveness. Of course, knowing that there is an impact of time on these interventions is likely not that useful for application, as we need recommendations for how long to use each specific pedagogy.

In order to answer the question of how long we should use each form of instruction, I analyzed the standardized test results for each type of intervention for multiple different periods of time. However, I do want to point out four potential flaws in this analysis. First, as I previously mentioned, effect sizes tend to get lower on average than longer studies. Second, these effect sizes were based on a random effects analysis, so there is some contamination in the instruction variables. Third, while researchers tend to use concentrated blocks of instruction, teachers might not. For example, we might find a study in which the researchers had students practice writing summaries every day for a month. However, a teacher might instead have students practice writing summaries once a week for a year. Results for the former might not reflect results of the latter. Fourth, most strategy instruction studies examine 1-2 strategies at a time, so the results below might not be reflective of instructional practices in which teachers taught new strategies each month. Lastly, only one study lasted longer than a year, and it showed no significant effects.

Studies of 2 Weeks or Less in Duration:

Instruction Type	Number of Effects	Mean ES	95% CI
Vocabulary	2	.35	[-1.61, 2.32]
Strategy	4	.35	[-.77, 1.48]
Cognitive Strategies/Skill	5	.77	[-0.58, 2.12]
Content/Background Knowledge	2	.15	[-0.79, 1.10]
Meta-Cognition	2	.44	[-4.26, 5.14]

Table 6: *Results for Studies Lasting 2 Weeks or Less*

Studies of 3-6 Weeks in Duration:

Instruction Type	Number of Effects	Mean ES	95% CI
Vocabulary	1	.14	
Strategy	2	.40	[.21, .59]
Cognitive Strategies/Skill	3	.31	[-.06, .69]
Graphic Organizers	3	.16	[-87, 1.20]
Meta-Cognition	1	.42	
Morphology	1	.14	
Technology	1	.14	

Table 7: *Results for Studies Lasting 3-6 Weeks*

Studies of 7-12 Weeks in Duration:

Instruction Type	Number of Effects	Mean ES	95% CI
Vocabulary	3	.54	[-.66, 1.65]
Strategy	15	.05	[-.13, .24]
Cognitive Strategies/Skill	12	.11	[-.17, .39]
Reciprocal Teaching	4	.45	[.17, .73]
Graphic Organizers	4	.16	[-.38, .70]
Content/Background Knowledge	1	.03	
Meta-Cognition	5	.34	[.07, .61]
Technology	1	.45	

Table 8: *Results for Studies Lasting 7-12 Weeks*

Studies of 13-35 Weeks in Duration

Instruction Type	Number of Effects	Mean ES	95% CI
Vocabulary	7	.22	[-.06, 51]
Strategy	9	.02	[-.09, .14]
Cognitive Strategies/Skill	3	.13	[-.14, .40]
Reciprocal Teaching	1	.25	
Graphic Organizers	1	.25	
Content/Background Knowledge	4	.19	[-.04, .43]
Meta-Cognition	4	.08	[-.12, .29]
Technology	1	.09	
Morphology	3	0	[-.26, .26]

Table 9: *Results for Studies Lasting 13-35 Weeks*

Studies of 1 Year or Longer in Duration

Instruction Type	Number of Effects	Mean ES	95% CI
Vocabulary	4	.14	[-.20, .48]
Cognitive Strategies/Skill	1	.46	
Graphic Organizers	1	.46	
Content/Background Knowledge	2	.24	[-2.48, 2.97]
Meta-Cognition	1	.46	

Table 10: *Results for Studies Lasting 1 Year or Longer*

The above results seem to support the common hypothesis that strategy instruction shows diminishing effects over time, and that background knowledge instruction requires longer time periods to show benefits. This research does not lend support to the hypothesis that vocabulary instruction requires longer durations to be effective. Indeed, vocabulary instruction showed the highest benefit between 13-35 weeks. That said, I think we need to make some specific caveats to these results.

First, these results seem to suggest that longer durations are beneficial for content instruction, because the only significant effect for content instruction was found at the one-year mark. However, only two studies examined the impact of background knowledge in this time frame, and strong interpretations should not be made based on these two studies.

Moreover, of these two studies, only one study showed significant effects. Swanson 2017, looked at the impact of vocabulary and content knowledge instruction and showed negligible results. The second study (Solis 2015) looked at the impact of combining content instruction, vocabulary instruction, cognitive strategies, meta-cognition strategies, and graphic organizers on reading comprehension results. The Solis study showed a mean effect size of .46. This therefore means that we only found evidence for a statistically significant effect of content instruction on reading comprehension results when the instruction was combined with strategy instruction. Second, these results seem to suggest the strongest

benefit for strategy instruction appeared to be at the two-week or less mark. Third, as most strategy instruction studies only looked at a small number of strategies at a time, we can only assume that these results show diminishing returns for specific strategies, not strategies in general.

It would be reasonable for teachers to look at this data and decide that they should not teach a specific strategy for more than two weeks of a year. However, it would not make sense for teachers to look at this data and decide that strategy instruction in totality should last only two weeks. Lastly, the ability to write a good summary or identify abstract ideas in a text is not so much of a strategy, as it is an important comprehension skill that students need to practice every year. Therefore, the below dose recommendations should be viewed on a per-year basis.

To make this information more digestible for teachers, I have created two charts below. The first chart shows the effective range of dosage for each intervention type as found in the above research. In this context, *effective* is defined as statistically significant. The second chart shows the optimal range of dosage. In this context *optimal* is defined as the highest effect size.

Effective Range:

Instructional Type	**Dose**
Vocabulary	7-35 weeks
Cognitive Skills	6 weeks or less
Meta-Cognition	2 -12 weeks
Graphic Organizers	13 weeks or more
Reciprocal Teaching	7-35 weeks
Content Knowledge	1 Year or Longer

Table 11: *The Effective Range of Dose for Comprehension Pedagogies*

Optimal Range:

Instructional Type	Dose
Vocabulary	7-12 weeks
Cognitive Skills	2 weeks or less
Meta-Cognition	2 weeks or less
Graphic Organizers	1 year or more
Reciprocal Teaching	7-12 weeks
Content Knowledge	1 Year or Longer

Table 12: *The Optimal Range of Dose for Comprehension Pedagogies*

Two longitudinal studies on the effects of knowledge building curriculum on reading comprehension scores were published, after the completion of my meta-analysis was done. One study was by Grissmer, et al and one study was by Kim, et al. Because I was not able to include these studies in my meta-analysis, I have included a review of each study below.

Grissmer, et al. published a four-year-long experimental working paper with eight other scholars. These researchers looked at the impact of a knowledge building intervention on standardized English language arts outcomes. The study employed an RCT study design, had a sample size of 6652, and used standardized, norm referenced testing. In this study, students entered into lotteries to try and attend highly competitive charter schools, which also used knowledge building curriculums. The researchers then compared the students who entered the lottery and won, vs the students who entered, but did not win and had to attend public school education. The study did have a high attrition rate of 34.1%; however, high attrition rates are often quite normal for longer studies. The students in the treatment group showed a mean effect size for English proficiency, of .47. This would suggest that the study results showed a strong positive result on English proficiency. The study also showed particularly positive

benefits for low-income students with an effect size of 1.29 on English proficiency.

That said, I do have some caveats that I would like to make in regard to this study. First, (to the best of my knowledge) this study did not specifically measure reading comprehension, which means it would not have met the inclusion criteria I used or established for my previous meta-analysis of this topic. Instead, this study used an assessment that gave a general score for literacy achievement. Two thirds of the marks for the assessment were based on comprehension ability and one third was based on writing ability.

Second, within this study design, students in the treatment group attended highly competitive charter schools. Indeed, these schools were so competitive that parents entered into lottery pools for their students to attend. Conversely, students in the comparison group stayed in the public school system and received business as usual instruction. This means that there is no real way to know what instruction looked like in the comparison group. Ostensibly, this is not really a study on the effects of knowledge building on reading comprehension, but rather on the impacts of specific charter schools. In order to assume that the effect size here represents improvements in reading comprehension, caused by a knowledge building curriculum, we also have to assume that there were no other significant differences between these charter schools and the public schools. This assumption does seem to be quite a stretch. None of this is to say that this research is not important or that it does not build on the scientific understanding of the topic. However, it cannot be used as definitive proof that “knowledge building curriculums’ increase reading comprehension scores.

In the Kim, et al, study 2156 grade 1-2 students were randomly selected from multiple schools. In the treatment group students received instruction based on the Model of Reading Engagement (MORE), which used systematic knowledge building curriculum for ELA, social studies, and science. In the control group grade 2 students used the Expeditionary Learning program, which is also a systematic knowledge building program, for their ELA programming. However, all other instruction was business as usual. In the control group there was no district wide social studies or science curriculum. The researchers in this study used standardized assessments, specifically the MAP comprehension assessment, to measure reading comprehension results. The instruction in this study lasted one school year; however, delayed post test results were measured in the fall of the following year. Their results can be seen in the following chart.

Table 6

Descriptive Statistics for Analytic Sample (Means and Standard Deviations) of Study Variables and Pairwise Correlation Matrix

Variable	Treatment		Control		Correlation									
	M (*SD*)	*n*	*M* (*SD*)	*n*	1	2	3	4	5	6	7	8	9	10
Reading: Total MAP scores across time periods														
1. T0	167.73 (15.61)	1,176	171.98 (15.38)	980	—									
2. T1	173.47 (15.70)	1,168	177.91 (16.01)	960	.89	—								
3. T2	173.06 (16.27)	1,169	176.50 (16.51)	956	.80	.80	—							
Science content reading (T3)														
4. Overall	10.97 (4.41)	1,176	10.97 (4.04)	980	.60	.61	.65	—						
5. Near-transfer subtest	3.66 (1.75)	1,176	3.49 (1.58)	980	.42	.42	.45	.73	—					
6. Mid-transfer subtest	3.66 (1.93)	1,176	3.65 (1.81)	980	.50	.52	.54	.83	.40	—				
7. Far-transfer subtest	3.65 (1.83)	1,176	3.83 (1.80)	980	.50	.51	.55	.81	.36	.54	—			
Vocabulary knowledge depth (T3)														
8. Overall	33.27 (9.76)	1,176	34.23 (9.07)	980	.55	.57	.59	.57	.49	.43	.44	—		
9. Taught words	19.77 (5.82)	1,176	20.64 (5.47)	980	.52	.54	.55	.53	.45	.40	.42	.94	—	
10. Untaught words	13.50 (4.68)	1,176	13.59 (4.42)	980	.49	.51	.53	.53	.45	.40	.40	.90	.70	—

Note. Domain-general MAP reading percentile ranks for the treatment group at Time 0, Time 1, and Time 2 MAP. RIT scaled scores correspond to the 38th, 38th, 45th percentile ranks for the treatment group, and 49th, 49th, and 54th percentile ranks for the control group. MAP = Measure of Academic Progress, Reading Total (RIT) scaled score; T0 = Time 0 (January of Grade 1), T1 = Time 1 (April of Grade 1), T2 = Time 2 (Follow-up 1, September of Grade 2), T3 = Time 3 (Follow-up 2, March of Grade 2).

Table 13: *Kim 2023 Results*

To standardize these results, I calculated Cohen's d effect sizes by comparing the post test scores between groups. The results can be seen in the below chart.

Assessment	Treatment Mean (SD)	Treatment N	Control Mean (SD)	Control N	Effect Size
Map Pre-Test	167.73 (15.61)	1176	171.98 (15.38)	980	-0.27
Map Post Test	173.47 (15.70)	1168	177.91 (16.01)	960	-0.28
Map Delayed Post Test	173.06 (16.27)	1169	176.50 (16.51)	956	-0.21
Science Overall	3.66 (1.75)	1176	3.49 (1.58)	980	0
Science Far Transfer Test	3.65 (1.83)	1176	3.83 (1.80)	980	-0.09

Table 14: *Kim Results Tabulated with Cohen's d*

In the Kim, 2023 study, the treatment group did worse than the control group on almost every assessment. While this study has been cited as evidence that knowledge building curriculums will increase reading comprehension scores, the actual study results seem to suggest a negative impact. If we calculate the effect size based on gains, rather than based on post test differences, the control group does worse. However, the difference is negligible.

In both these studies one group receives knowledge building focused instruction and the other group receives business as usual instruction. This means that we cannot fully attribute the results of either study to the impact of knowledge building. Of course, herein lies one of the most challenging elements of education research. For a study to measure a fixed effect, we need to isolate the study variable and make sure only one variable is different between the control group and treatment group. This is easy to do in small studies with short time horizons. However, it is very difficult, if not impossible to do with large scale longitudinal studies like this one. This is why we need to look at multiple studies and studies with different designs to be sure that an intervention or pedagogy is effective. Caveats aside, I do think that these are both important and rigorous studies. However, only one of the two studies showed a positive effect size, when comparing the treatment group to the control group. The Grissmer study helps to start to build a body of scientific literature that background knowledge interventions can be effective over long-time

horizons. However, to quote Natalie Wexler's excellent and nuanced summary of the study: “One study can never be definitive, although the researchers tried to eliminate potential sources of “bias” that might limit the applicability of the findings.” I do hope to see more research on this topic moving forward.

Written by Nathaniel Hansford
Contributed to by Rachel Shechter
Last Edited 2023-04-20

References:

-Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading Comprehension Instruction for Students With Learning Disabilities, 1995—2006: A Meta-Analysis. Remedial and Special Education, 31(6), 423–436. https://doi.org/10.1177/0741932509355988

-Donegan RE, Wanzek J. Effects of Reading Interventions Implemented for Upper Elementary Struggling Readers: A Look at Recent Research. Read Writ. 2021 Oct;34(8):1943-1977. doi: 10.1007/s11145-021-10123-y. Epub 2021 Jan 26. PMID: 34720414; PMCID: PMC8553009.

-Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. Journal of Educational Psychology, 92(4), 605-619. doi:https://doi.org/10.1037/0022-0663.92.4.605

-Fukkink, Ruben & Glopper, Kees. (1998). Effects of Instruction in Deriving Word Meaning from Context: A Meta-Analysis. Review of Educational Research - REV EDUC RES. 68. 450-469. 10.2307/1170735.

-Gajria M, Jitendra AK, Sood S, Sacks G. Improving comprehension of expository text in students with LD: a research synthesis. J Learn Disabil. 2007 May-Jun;40(3):210-25. doi: 10.1177/00222194070400030301. PMID: 17518214.

-Grissmer, D,. Buddin, R,. Berends, M,. Willingham, D,. DeCoster, J,. Duran, C,. Hulleman, C,. Murrah & Evans, T. (2023). A Kindergarten Lottery Evaluation of Core Knowledge Charter Schools: Should Building General Knowledge Have a Central Role in Educational and Social Science Research and Policy? *Annenberg University.* https://edworkingpapers.com/sites/default/files/ai23-755.pdf

-Hansford, N. (2023). What's Better for Reading Comprehension: Strategy or Content Knowledge Instruction? Teaching by Science. https://www.teachingbyscience.com/reading-comprehension

Kim, James & Burkhauser, Mary & Relyea, Jackie & Gilbert, Josh & Scherer, Ethan & Fitzgerald, Jill & Mosher, Douglas & McIntyre, Joseph. (2022). A Longitudinal Randomized Trial of a Sustained Content Literacy Intervention From First to Second Grade: Transfer Effects on Students' Reading Comprehension. Journal of Educational Psychology. 115. 10.1037/edu0000751.

-Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching Students to Generate Questions: A Review of the Intervention Studies. Review of Educational Research, 66(2), 181–221. https://doi.org/10.2307/1170607

-Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. Developmental Psychology, 46(6), 1556–1579. https://doi.org/10.1037/a0020612

-Wexler, N. (2023). Dramatic New Evidence That Building Knowledge Can Boost Comprehension And Close Gaps. Forbes. https://www.forbes.com/sites/nataliewexler/2023/04/09/dramatic-new-evidence-that-building-knowledge-can-boost-comprehension-and-close-gaps/amp/

-Talbott, E., Lloyd, J. W., & Tankersley, M. (1994). Effects of Reading Comprehension Interventions for Students with Learning Disabilities. *Learning Disability Quarterly*, *17*(3), 223–232. https://doi.org/10.2307/1511075