

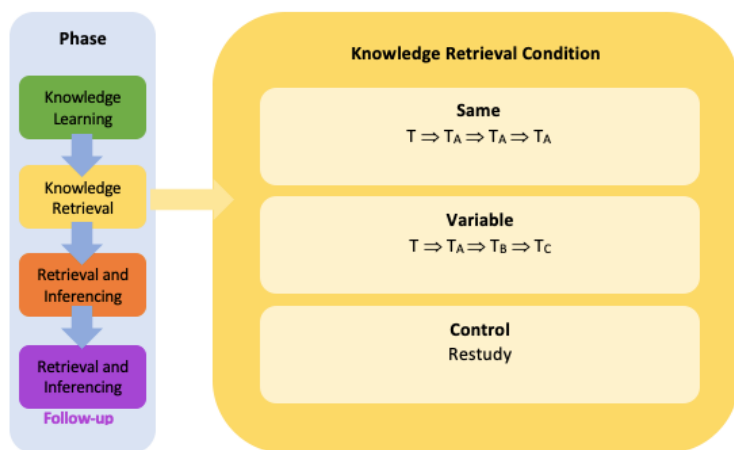
STUDY PROTOCOL

OFFICIAL TITLE: The role of knowledge retrieval in inference-making among struggling readers

NCT NUMBER:

DOCUMENT DATE: 20 December 2022

Overview of Aim 1 Research Strategy and Design: Role of Knowledge Retrieval.



A randomized controlled trial design will establish the effectiveness of two knowledge retrieval techniques versus a control condition (See Figure 2). This research design builds on Butler (2010) to investigate how conditions of initial learning affect retention of newly acquired knowledge and its retrieval to form inferences that draw on that knowledge. Using the same general procedure, students will be taught a knowledge base comprised of 18 facts about Egypt (i.e., Knowledge Learning). Second, after learning the knowledge base to criterion (perfect recall), students will practice retrieving the knowledge base

across four testing sessions spaced 24 hours apart (i.e., Knowledge Retrieval). Depending on the condition to which students get randomized, they will receive the same questions across testing sessions, variable questions, or, in the control condition, will re-study the knowledge base. Next, students will read 6 passages about Egypt and then answer literal and inferential questions about the passages. The Knowledge Retrieval and Inferencing Phase provides students the opportunity to retrieve the knowledge base to form inferences while reading text. Immediately after reading the passages, students will get re-tested on their retention of the knowledge base and use of the knowledge base to form inferences. They will also get retested one week and one month later (i.e., Follow up Phase) to examine students' long-term retention of the knowledge base and use to make inferences. **We hypothesize that the Variable condition will be most effective. Variable encoding is thought to result in improved knowledge retention due to an increased number of potential retrieval routes; thereby increasing the probability that the knowledge will be retrieved and integrated to form inferences.**

Materials and Description of Knowledge Retrieval Conditions. The materials for this experiment include 6 passages selected from *Egypt World* (Caldwell, 2013), an expository text with a Lexile level of 1160. The PI will select three facts/concepts (18 total concepts) from the passages. Concepts will represent the knowledge base that students learn in the Knowledge Learning phase. Once students learn the 18 concepts to perfect recall, they will practice retrieving the knowledge in the Knowledge Retrieval phase. Students will get randomized to one of three conditions: Same, Variable, or Control conditions. In the Same Condition, students will repeatedly retrieve the 18 concepts using the same test over time. Questions will not facilitate rote recall of the knowledge base. Rather, questions will facilitate elaboration of the knowledge base facts. Randomly ordering questions will counterbalance forms across participants. In the Variable Condition, students will receive a different test each session with new and different questions that facilitate elaboration of the knowledge base concepts. In the Control condition, student will re-study the knowledge base items with items depicted in colorful pictures. Students will look at the picture, state the concept aloud, but receive no feedback for correct or incorrect responses.

Procedure. The experiment will occur on the computer using E-Prime software with all stimuli presented orally via headphones, a procedure successfully used in our previously studies. The computer will capture retrieval responses to score accuracy and reaction time. The experiment lasts 6 weeks. In week 1 and in a single session, students will learn the knowledge base to perfect mastery. In week 2, students will practice retrieving the knowledge base for 10 minutes each day from Monday-Thursday. On Friday, students will read the 6 passages and answer inference questions that draw on the knowledge base. Reassessment of students' retention of the knowledge base and ability to form inferences with the knowledge base will occur 1-week and one-month later.

Sample. We will recruit a total of 158 demographically diverse middle grade readers and English Learners, all of whom perform below grade level benchmarks in reading as measured by the Fastbridge Reading Fluency Assessment, Fastbridge aReading Assessment, or Iowa Statewide Assessment of Student Progress. Due to task demands, students diagnosed with a significant cognitive impairment, participating in the Life Skills class (e.g., severe autism, behavior, or cognitive disability) or Level 1 English for Newcomers class will be excluded.

Inclusion and Exclusion of Participants. Students will be screened using the Test of Silent Reading Efficiency and Comprehension. Students performing below a standard score of 93 will qualify for participation in the study.

Students with a clinical diagnosis of significant cognitive disability or significant behavioral disability will be excluded.

Assessment Battery. Participants will also complete a comprehensive assessment battery to identify individual difference factors that may impact knowledge retrieval, inferencing, and relationships among knowledge retrieval and reading-related constructs. Assessments will measure the constructs of reading comprehension, inferencing, working reading efficiency, metacognition, working memory, and background knowledge.

- **Knowledge Retrieval and Relation to Inferencing.** The Experimental Knowledge task is an individually administered experimental procedure delivered on the computer. It teaches new knowledge then measures the accuracy and efficiency by which the individual can retrieve that knowledge. After learning the knowledge base, students read passages and form inferences with the accuracy and efficiency of inferencing measured.
- **Reading Comprehension** will be measured by the *Gates-MacGinitie Reading Test-4th Edition* (GMRT-4; MacGinitie et al., 2000), a standardized, group-administered, multiple-choice assessment of reading comprehension with internal consistency for students in grades 5-8 that exceeds 0.80, and *Iowa Assessment for Reading* (Iowa Assessment Program, 2015), a group administered standardized measure of reading comprehension with internal consistency for students in grades 5-8 above 0.80. The *Weschler Individual Achievement Text-III* (WIAT-III; Weschler, 2009) is an individually administered, standardized assessment of reading comprehension with internal consistency for students in grades 5-8 exceeding 0.90. *Qualitative Reading Inventory-5* (QRI-5; Leslie & Caldwell, 2011) is an informal measure of students' recall of text and understanding of text when prompted by questions. Students will be administered 2 sixth grade level passages on ancient Egypt (i.e., *Nile* and *Building Pyramids* with a Lexile of 850). Internal consistency was 0.83 for students in the middle grades (Barth & Elleman, 2017).
- **Inferencing.** *Test of Language Competence-Expanded Inference Task* asks students to read three sentences and form an inference. *Bridge-IT* (Barth et al., 2015) asks students to read a five-sentence text, form an inference, and determine if the next sentence provided continues the story. Internal consistency for students in the middle grades ranges from 0.83-0.87. The *Clinical Evaluation of Language Fundamentals-5* (CELF-5; Wiig et al., 2013), Metalinguistics, Making Inferences subtest, an individually administered assessment, measures students' ability to make gap-filling inferences on the basis of causal relationships or event chains depicted in short narratives that are presented both orally and in print form. The internal consistency exceeds 0.80 for students in the middle grades.
- **Word Reading Efficiency.** *Test of Word Reading Efficiency-2* (TOWRE-2; Torgesen et al., 2012) includes Sight Word Efficiency, a 45-second assessment of students' word reading accuracy and speed and Phonemic Decoding Efficiency, a 45-second assessment of students' non-word reading accuracy and speed. Alternate form reliability exceeds 0.90 for students in grades 5-8. *Fastbridge Reading fluency* (Christ, 2015) is an individually administered measure of oral reading fluency that is administered by area schools. Internal consistency exceeds 0.80. The *Test of Silent Reading Efficiency and Comprehension* (TOSREC; Wagner et al., 2010) is a brief, group administered measure of silent word reading and comprehension of connected text. TOSREC is both a measure of word reading efficiency and reading comprehension. Average alternate form reliability exceeds 0.80 for students in the middle grades. The *Test of Silent Contextual Reading Fluency-2nd Edition* (TOSCRF2: Hammill, Wiederholt, & Allen, 2014) is a brief, group administered measure of silent reading fluency and comprehension of connected text. Average alternate form reliability exceeds 0.80 for students in the middle grades. *CBM-Maze* is a brief, group based assessment of reading fluency and comprehension.
- **Metacognition.** *Metacognition survey* (Karpicke et al., 2009) asks students to report study strategies used when learning from text and includes questions about repeated studying and retrieval practice. The *Contextualized Reading Strategy Survey* (CReSS; Denton et al., 2014) is a 49-item self-report survey designed to assess a collection of learning and comprehension strategies that students might use when reading for school. CReSS items relate to four factors: evaluation and knowledge integration, pragmatic strategies, self-regulation, and help-seeking. Cronbach's coefficient alphas for middle grades are moderate.
- **Working Memory.** *Working Memory Test Battery for Children* (Gathercole et al., 2004) uses Word List Recall that asks the student to repeat a list of monosyllabic words, Non-word List Recall that asks the student to repeat a list of non-words and Listening Recall Test that requires that the student listen to a series of short sentences, judge their accuracy, and recall the final word of each sentence.
- **Background Knowledge.** *Kauffman Brief Intelligence Test-2* (KBIT-2; Kauffman & Kauffman, 2004). The

Verbal Knowledge subtest, a norm-referenced, individually administered assessment of expressive word and world knowledge, has an internal consistency that ranges from 0.89-0.94 for students in grades 5-8. *Gates MacGinitie Vocabulary Test* (2000) is a group administered test of receptive vocabulary that has an internal consistency that exceeds 0.80 for students in grades 5-8. *Egyptian Content Knowledge* is a researcher developed assessment of one's knowledge of Egypt that has an internal consistency that exceed 0.80 for students in the middle grades (Barth & Elleman, 2017).

Overview of Aim 2 Research Strategy and Design: Inference Intervention

Research Design. A randomized controlled trial design will establish the effectiveness of an intervention that enhances knowledge retrieval and explicitly teaches inference making compared to a control group. Intervention students will practice (a) retrieving newly acquired knowledge necessary to understand intervention texts; (b) using a strategy to identify and prompt integration of important text content with relevant prior knowledge; (c) using context clues and prior knowledge to define novel words in text; and (d) explaining how connections between text and knowledge helped them understand the text, character's perspectives, and author's purpose. Aim 1 will inform intervention component A; the studies described above previously tested intervention components B-D albeit for a shorter duration than what's proposed in Aim 2.

Intervention. Students randomized to the Inference Treatment Condition will receive intervention in small groups of 3-6 students, delivered in 40-minute sessions, 2-3 times per week for a total of 10 weeks. The intervention will pre-teach students a set of 20 essential facts about Egypt. Results from Aim 1 will identify the most effective knowledge retrieval technique to ensure that the students accurately encode the knowledge base in memory and can easily retrieve it while reading. Following knowledge retrieval practice, students will engage in expository and narrative texts about Egypt. Expository texts include passages from *Egypt World* (Caldwell, 2013) with a Lexile level of 1160 and *Mummies and Pyramids*, a 106-page text with a Lexile of 650. Students will also read the narrative text *Tut! Tut!* (Scieska, 2006), a 73-page book with a Lexile of 700. While reading these texts, students will practice using a text clue strategy to retrieve relevant knowledge from memory. After students retrieve the target term or vocabulary word, they will check it for accuracy against the text and discuss the target term with the tutor. Second, students will practice activating and integrating prior knowledge. Students will learn to make in-text connections, connections across the 3 intervention texts, and connections between text and knowledge. The tutor will ask students to justify these connections and evaluate how these connections enhance understanding. Tutor feedback will correct errors and scaffold connections. Third, students will practice understanding character perspectives and author purpose while reading the narrative text *Tut! Tut!* Students will learn how to make inferences about character motives and intent using text clues and connections to background knowledge. They will justify the importance of the information to the text and discuss how it enhances understanding. Fourth, students will answer inference questions that require using text and knowledge. A simple graphic organizer will support and scaffold inferencing. Finally, throughout the intervention, students will practice retrieving the knowledge base from memory using flash cards, completed in pairs with students providing each other feedback.

Sample. The study will randomly select a total of 158 demographically diverse middle grade struggling readers and English Learners for participation, with "struggling reader" defined as students who have performed below grade level benchmarks as established by the local school district. Measures could include reading fluency as measured by the Fastbridge Reading Fluency Assessment, reading comprehension as measured by the Fastbridge Reading Assessment or Iowa Statewide Assessment of Student Progress. Due to task demands, the study will exclude students diagnosed with a significant cognitive impairment, participating in the Life Skills class (e.g., severe autism, behavior, or cognitive disability), or participating in Level 1 English for Newcomers. Half of students will get randomized to the Inference Treatment condition and half to the Comparison condition delivered by the school district. Approximately 657 students in grades 5-8 (10-14 years of age) will comprise the potential subjects for the study. Students are 25% minority-non-Hispanic, 54% Hispanic, and 21% White; 75% received reduced lunch; 16% live below the poverty line; 14% are in special education; and 51% are male.

Assessment Battery. Participants will also complete a comprehensive assessment battery to measure intervention effectiveness and identify individual difference factors that may impact response to intervention. Outcome assessments will measure the constructs of inference making and reading comprehension. Additional assessments will measure the constructs of working reading efficiency, cognitive load, metacognition, working memory, background knowledge, and knowledge retrieval.

- **Knowledge Retrieval and Relation to Inferencing.** The Experimental Knowledge task is an individually administered experimental procedure delivered on the computer. It teaches new knowledge then measures the accuracy and efficiency by which the individual can retrieve that knowledge. After learning the knowledge base, students read passages and form inferences with the accuracy and efficiency of inferencing measured.
- **Reading Comprehension** will be measured by the *Gates-MacGinitie Reading Test-4th Edition* (GMRT-4; MacGinitie et al., 2000), a standardized, group-administered, multiple-choice assessment of reading comprehension with internal consistency for students in grades 5-8 that exceeds 0.80, and *Iowa Assessment for Reading* (Iowa Assessment Program, 2015), a group administered standardized measure of reading comprehension with internal consistency for students in grades 5-8 above 0.80. The *Weschler Individual Achievement Text-III* (WIAT-III; Weschler, 2009) is an individually administered, standardized assessment of reading comprehension with internal consistency for students in grades 5-8 exceeding 0.90. *Qualitative Reading Inventory-5* (QRI-5; Leslie & Caldwell, 2011) is an informal measure of students' recall of text and understanding of text when prompted by questions. Students will be administered 2 sixth grade level passages on ancient Egypt (i.e., *Nile* and *Building Pyramids* with a Lexile of 850). Internal consistency was 0.83 for students in the middle grades (Barth & Elleman, 2017).
- **Inferencing.** *Test of Language Competence-Expanded* Inference Task asks students to read three sentences and form an inference. *Bridge-IT* (Barth et al., 2015) asks students to read a five-sentence text, form an inference, and determine if the next sentence provided continues the story. Internal consistency for students in the middle grades ranges from 0.83-0.87. The *Clinical Evaluation of Language Fundamentals-5* (CELF-5; Wiig et al., 2013), Metalinguistics, Making Inferences subtest, an individually administered assessment, measures students' ability to make gap-filling inferences on the basis of causal relationships or event chains depicted in short narratives that are presented both orally and in print form. The internal consistency exceeds 0.80 for students in the middle grades.
- **Word Reading Efficiency.** *Test of Word Reading Efficiency-2* (TOWRE-2; Torgesen et al., 2012) includes Sight Word Efficiency, a 45-second assessment of students' word reading accuracy and speed and Phonemic Decoding Efficiency, a 45-second assessment of students' non-word reading accuracy and speed. Alternate form reliability exceeds 0.90 for students in grades 5-8. Fastbridge Reading fluency (Christ, 2015) is an individually administered measure of oral reading fluency that is administered by area schools. Internal consistency exceeds 0.80. The Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner et al., 2010) is a brief, group administered measure of silent word reading and comprehension of connected text. TOSREC is both a measure of word reading efficiency and reading comprehension. Average alternate form reliability exceeds 0.80 for students in the middle grades.
- **Metacognition.** *Metacognition survey* (Karpicke et al., 2009) asks students to report study strategies used when learning from text and includes questions about repeated studying and retrieval practice. The Contextualized Reading Strategy Survey (CReSS; Denton et al., 2014) is a 49-item self-report survey designed to assess a collection of learning and comprehension strategies that students might use when reading for school. CReSS items relate to four factors: evaluation and knowledge integration, pragmatic strategies, self-regulation, and help-seeking. Cronbach's coefficient alphas for middle grades are moderate.
- **Working Memory.** *Working Memory Test Battery for Children* (Gathercole et al., 2004) uses Word List Recall that asks the student to repeat a list of monosyllabic words, Non-word List Recall that asks the student to repeat a list of non-words and Listening Recall Test that requires that the student listen to a series of short sentences, judge their accuracy, and recall the final word of each sentence.
- **Background Knowledge.** *Kauffman Brief Intelligence Test-2* (KBIT-2; Kauffman & Kauffman, 2004). The Verbal Knowledge subtest, a norm-referenced, individually administered assessment of expressive word and world knowledge, has an internal consistency that ranges from 0.89-0.94 for students in grades 5-8. *Gates MacGinitie Vocabulary Test* (2000) is a group administered test of receptive vocabulary that has an internal consistency that exceeds 0.80 for students in grades 5-8. *Egyptian Content Knowledge* is a researcher developed assessment of one's knowledge of Egypt that has an internal consistency that exceed 0.80 for

students in the middle grades (Barth & Elleman, 2017).